

ULTRA-COMPACT HIGH VELOCITY CLOUDS IN THE ALFALFA HI SURVEY: CANDIDATE LOCAL GROUP GALAXIES?

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Elizabeth Ann Kovenz Adams

January 2014

© 2014 Elizabeth Ann Kovenz Adams

ALL RIGHTS RESERVED

ULTRA-COMPACT HIGH VELOCITY CLOUDS IN THE ALFALFA HI SURVEY: CANDIDATE LOCAL GROUP GALAXIES?

Elizabeth Ann Kovenz Adams, Ph.D.

Cornell University 2014

The increased sensitivity and spatial resolution of the ALFALFA HI survey has resulted in the detection of ultra-compact high velocity clouds (UCHVCs). These objects are good candidates to represent low mass gas-rich galaxies in the Local Group and Local Volume with stellar populations that are too faint to be detected in extant optical surveys. This idea is referred to as the “minihalo hypothesis”.

We identify the UCHVCs within the ALFALFA dataset via the use of a 3D matched filtering signal identification algorithm. UCHVCs are selected based on a compact size ($< 30'$), separation from Galactic HI ($|v_{LSR}| > 120 \text{ km s}^{-1}$) and isolation. Within the 40% complete ALFALFA survey ($\alpha.40$), 59 UCHVCs are identified; 19 are in a most-isolated subset and are the best galaxy candidates.

Due to the presence of large HVC complexes in the fall sky, most notably the Magellanic Stream, the association of UCHVCs with existing structure cannot be ruled out. In the spring sky, the spatial and kinematic distribution of the UCHVCs is consistent with simulations of dark matter halos within the Local Group. In addition, the HI properties of the UCHVCs (if placed at 1 Mpc) are consistent with both theoretical and observational predictions for low mass gas-rich galaxies. Importantly, the HI properties of the UCHVCs are consistent with those of two recently discovered low mass gas-rich galaxies in the Local Group and Local Volume, Leo T and Leo P.

Detailed follow-up observations are key for addressing the minihalo hypothe-

sis. High resolution HI observations can constrain the environment of a UCHVC and offer evidence for a hosting dark matter halo through evidence of rotation support and comparison to theoretical models. Observations of one UCHVC at high resolution (15'') reveal the presence of a clumpy HI distribution, similar to both low mass galaxies and circumgalactic compact HVCs. An extended envelope containing $\sim 50\%$ of the HI flux is resolved out by the array configuration; observations at lower spatial resolution can recover this envelope and constrain the overall morphology and environment.

The most direct way to address the minihalo hypothesis is by detection of a stellar counterpart, immediately identifying a UCHVC as a galaxy and allowing a distance to be measured. We have selected a sample of best galaxy candidates from the UCHVCs based on isolation, compactness, surface brightness, and kinematics. We are undertaking targeted optical observations of these systems in two filters to conduct a focused search for a coherent stellar population. Observations are in-hand for 29 systems, and an analysis of a single system is presented as a test case. These data were obtained via "shared-risk" observing, and analysis for all systems is awaiting further pipeline development.

If (some of) the UCHVCs represent nearby low mass galaxies, they will help us understand the evolution of the lowest mass galaxies and address the small scale crisis in cosmology. Understanding the nature of the UCHVCs is a complicated and ongoing project. Both optical and HI synthesis imaging data will continue to be acquired and analyzed in order to address the minihalo hypothesis. Future HI surveys of nearby galaxy groups will be able to robustly address the minihalo hypothesis by being sensitive to UCHVCs in other galaxy groups.

BIOGRAPHICAL SKETCH

Elizabeth Ann Kovenz Adams is proud of having two middle names, prefers to be called Betsey, and is a Hoosier through and through. She grew up in La Porte, Indiana at the confluence of three different time zones. She attended Indiana University in Bloomington, IN where she studied astronomy, math, hapkido and IU basketball. She spent a year abroad in Australia where she studied upside-down constellations, variable red giants, and learned to rock climb. Upon her return to Bloomington, she completed two more years of course work and studied dwarf irregular galaxies in her spare time. After completing her degrees in astronomy and mathematics and earning a black belt in hapkido, she headed to Ithaca, NY for graduate school in astronomy at Cornell University. Her time in Ithaca was mostly spent searching for previously undiscovered low mass galaxies but she also discovered the joys and pains of long distance running. She is heading to AS-TRON in the Netherlands to continue her search for new galaxies and is eagerly anticipating the next adventure.

To the amazing women in my life who have served as role models and
inspiration.

ACKNOWLEDGEMENTS

Martha and Riccardo – I don't have the words to adequately thank you for the last six years. The support and encouragement you have offered me is phenomenal. Thank you for pushing me to grow and giving me a challenging open-ended project. I look forward to ALFALFA East and the continuation of this work.

The ALFALFA collaboration and the Undergraduate ALFALFA Team – It has been a joy to be a member of the ALFALFA team, and that's due in large part to the wonderful people who comprise the team. Thanks to everyone who helped produce the final ALFALFA data that I used in this dissertation – I never could have done it on my own. Becky – Thanks for inviting me to the UAT workshop so many times. I loved every sleep-deprived minute of it. John Cannon – Thank you for all your help and all the letters of recommendation. To my fellow denizens of Camuy Cave – Thank you for all the help over the years, especially answering my random IDL questions.

The NSF Graduate Research Fellowship, Brinson Foundation, NASA, and Cornell Graduate School offered me support in my research and travels.

Liese – Many, many thanks for all the support you have offered me over the years, professionally and personally. We'll eventually publish that paper and maybe one day you'll even let me buy you lunch.

Dave and Ira – Thank you for being on my committee, reading my dissertation, and always being interested in my work.

Sabrina and Ann – I can't imagine how my first few years of graduate school would have gone without you two in the Cave – it definitely hasn't felt the same since you left. Thank you for the epic 5 B's tour across America, TV nights, introducing me to feminist blogs, and only mocking my lack of any pop culture knowledge occasionally. Sadly, I believe Little Umi is lost for good.

Mom, Dad, Sam, Kelsey and the rest of my family – You are the ones who have always believed in me and encouraged me to follow my dreams. Thank you for your support and love over the years.

Betsy, Le Mai, Keith, Le Anh – It's been six years since we were all together in Bloomington, and it both feels like that was yesterday and a lifetime of ago. Wherever I may go and whatever I may end up doing, I know you will always be there for me and get me in ways that no one else can. Betsy – Well, we still can't manage to end up anywhere near each other. I think the Universe is still trying to recover from the fun and trouble we caused during the single year we lived together. May postdocs be easier than graduate school, and, if not, may we at least have each other to complain to. Keith – I've never found anyone that I can be as hilariously inappropriate with as you, and I doubt I ever will. Don't ever change. Le Mai – Save a spot for me on your homestead, and I'll join you one day. Until then, I'll think of you every time I do a puzzle. Le Anh – Other than being her amazing self, the best thing your sister ever did was introduce me to you. Thank you for hosting me on so many of my early trips back to Bloomington.

Carl, Drew, Babs, Manolis, and the rest of the Cornell astro grads – Graduate school would not have been nearly as much fun without such a great group of people to share with it. I'm going to miss beers at the Barn, picnics at Stewart Park, and the rest of our socializing.

The Hyers House – I can't imagine a better set of housemates. Thanks for the introduction to home brewing, house hikes, and all the rest.

Caty and Peter – Thank you for the undergraduate research experiences and your support during my applications to graduate school.

The Alexanders – From elementary school to middle school you fostered my love of science. Thank you.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vii
List of Tables	xi
List of Figures	xii
1 Introduction	1
1.1 The Small Scale Crisis	2
1.2 Solving the Problem	3
1.2.1 Relative Velocity of Dark Matter and Baryons	4
1.2.2 Reionization	4
1.2.3 Star Formation Feedback	5
1.3 Constraining the Problem by Finding New Galaxies	6
1.3.1 Ultra-faint Dwarf Galaxies	7
1.3.2 Leo T: A Gas-Rich Ultra-faint Dwarf	8
1.4 High Velocity Clouds as Dark Galaxies	9
1.5 Ultra-compact High Velocity Clouds as Gas-Rich UFDs	10
1.5.1 Ultra-compact High Velocity Clouds in ALFALFA	11
1.5.2 Proof of Concept: Leo P	11
1.6 Overview of the Thesis	12
2 Automated Signal Identification of Ultra-compact High Velocity Clouds in the ALFALFA HI Survey	14
2.1 ALFALFA, Source Identification, and UCHVCs	14
2.1.1 The ALFALFA HI Data	15
2.1.2 Source Identification in the ALFALFA Dataset	15
2.1.3 Motivation for a Specialized Algorithm	16
2.2 Methodology of the Algorithm	18
2.2.1 Matched Filtering	18
2.3 Implementation	21
2.3.1 The Templates	21
2.3.2 Galactic Hydrogen	22
2.3.3 Detection of Galaxies	23
2.3.4 Identifying a Source	23
2.3.5 Measuring Source Properties	25
2.4 Creation of Test Data	25
2.4.1 Simulating Artificial Sources	25
2.5 The Completeness of the Algorithm	26
2.6 Comparison to the Standard ALFALFA Scheme	30
2.7 Comparison to Visual Inspection	33

2.8	Accuracy of Reported Source Values	35
2.9	The Effects of the Isolation Criteria	39
2.10	Summary	47
3	Catalog of Ultra-compact High Velocity Clouds within the α.40 Footprint	48
3.1	Data	48
3.1.1	The ALFALFA Survey	50
3.1.2	Source Identification	50
3.1.3	Criteria for UCHVC Identification	51
3.1.4	Isolation	53
3.2	Catalog	59
3.2.1	Presentation of Catalog	59
3.2.2	Comparison to G10	69
3.2.3	Properties of the UCHVCs	69
3.2.4	Inferred Cloud Parameters	72
3.3	The UCHVCs as a Distinct Population	78
3.3.1	The UCHVCs in the Context of Large HVC Complexes . . .	80
3.3.2	UCHVCs in the Context of Galactic Halo Clouds	83
3.3.3	UCHVCs in the Context of the Small Cloud Population of GALFA-HI	86
3.4	UCHVCs as Minihalo Candidates	91
3.4.1	Previous Searches for Minihalos	94
3.4.2	Known Minihalos in the LG	98
3.4.3	Evidence for the UCHVCs as Minihalo Candidates	99
3.4.4	The UCHVCs as Galaxies	104
3.5	Conclusion	107
4	High Spatial Resolution Studies of the HI Distribution of Ultra-compact High Velocity Clouds	109
4.1	The Importance of the HI Distribution	109
4.1.1	Evidence for Interaction	110
4.1.2	Evidence for Rotation	110
4.1.3	Comparison to Models	111
4.2	A Case Study: AGC268069	112
4.3	The Observations	114
4.4	Data Preparation	115
4.4.1	Calibration	116
4.4.2	Dealing with Overlapping Passbands	117
4.5	Imaging	119
4.5.1	Robustness and Tapering	121
4.5.2	Imaging AGC268069	123
4.5.3	Flux Recovery	127
4.5.4	The Velocity Field	131

4.6	Discussion	132
4.6.1	HI Morphology: Three Cores and the Northern Extension	132
4.6.2	The Nature of AGC268069	138
4.7	Improving Constraints	139
4.8	Summary	140
5	Optical Observations of Ultra-compact High Velocity Clouds	141
5.1	Target Selection	141
5.1.1	The Sky Searched	142
5.1.2	Selecting the Best Galaxy Candidates	142
5.1.3	The Targets	147
5.2	The Observations	150
5.3	Data Reduction and Calibration	153
5.3.1	pODI: The “partially-filled” One Degree Imager	153
5.3.2	The Standard Calibration Procedure	154
5.3.3	Pipeline Processing	154
5.4	Post-Pipeline Processing	156
5.4.1	Photometric Calibration	157
5.5	Analysis of the Stellar Population	158
5.5.1	Identify and Measure Stars	158
5.5.2	Filtering the Stars	159
5.5.3	Smoothing the Filtered Population	160
5.6	Discussion	162
5.7	Summary	165
6	Conclusion and Future Directions	166
6.1	Identification of UCHVCs	167
6.2	The UCHVCs as a Class	167
6.3	Case Studies of UCHVCs	168
6.3.1	High Spatial Resolution HI Studies	169
6.3.2	Optical Searches for Stellar Counterparts	170
6.4	Future Work	170
6.4.1	Expanding the UCHVC Catalog	171
6.4.2	Optical Studies	171
6.4.3	High Resolution HI Imaging	173
6.4.4	Future HI Surveys	174
6.5	Summary	175
A	Source Code for the 3D Extraction Algorithm for UCHVCs	177

B	Detailed Reduction and Imaging Steps for VLA Data	190
B.1	Calibration	190
B.1.1	Flagging	191
B.1.2	Split the Data	192
B.1.3	Set the Flux Scale	193
B.1.4	Bandpass Calibration	193
B.1.5	Gain Calibration	195
B.1.6	Get the Flux Scale	197
B.1.7	Apply the Calibration	198
B.1.8	Check Calibration and Redo as Necessary	199
B.1.9	Split and “cvel” the Data	200
B.1.10	Continuum Subtraction	201
B.2	Imaging	203
B.2.1	Obtaining Clean Image Cubes	203
B.2.2	Creating a Mask	204
B.2.3	Pedestal Effect	208
B.2.4	Final Products: Moment Maps and Spectra	209
B.3	Image Testing	211
B.3.1	Combining SPWs	212
B.3.2	Robustness Levels	213
C	Detailed Reduction and Analysis Steps for WIYN/pODI Optical Data	216
C.1	Post-pipeline Processing	217
C.1.1	Cosmic Ray Removal	217
C.1.2	Shifting the Images	221
C.2	Photometric Calibration	224
C.2.1	Obtaining “Standard” Stars	224
C.2.2	Aperture Photometry of the Standard Stars	225
C.2.3	Getting Photometric Solutions	229
C.3	Photometry of the Stars	230
C.3.1	Locate the Stars	230
C.3.2	Aperture Photometry	235
C.3.3	Make the PSF	237
C.3.4	Do the PSF-Fitting Photometry	240
C.3.5	Aperture Correction	241
C.3.6	Transform Magnitudes to Standard System	243
C.4	Color Magnitude Diagram Analysis	244

LIST OF TABLES

3.1	ALFALFA UCHVCs in the $\alpha 40$ Survey	67
3.1	ALFALFA UCHVCs in the $\alpha 40$ Survey	68
3.1	ALFALFA UCHVCs in the $\alpha 40$ Survey	70
3.2	UCHVCs from G10 that Fail UCHVC Criteria	71
3.3	Inferred Cloud Properties	76
3.3	Inferred Cloud Properties	77
3.3	Inferred Cloud Properties	78
3.4	UCHVCs within $D = 25^\circ$ of a WvW complex	81
3.5	HI Content in the LG - HVCs and Galaxies	96
4.1	Summary of VLA/C Observations of AGC268069	115
4.2	Observational Set-up	115
4.3	Properties of VLA Data Cubes with Different Imaging Schemes . .	123
4.4	Properties of the HI Cores and Envelope	134
5.1	ALFALFA UCHVCs Targeted for Optical Observations	149
5.1	ALFALFA UCHVCs Targeted for Optical Observations	151
5.2	ALFALFA UCHVCs Observed in March WIYN Run	152
5.3	ALFALFA UCHVCs Observed in the April WIYN Run	152
5.4	ALFALFA UCHVCs Observed in the April CTIO Run	153

LIST OF FIGURES

2.1	Distribution of HI angular diameters and HI linewidths for UCHVCs, α .40 HVCs, and α .40 extragalactic detections.	17
2.2	Fraction of simulated sources detected by the signal extractor presented here.	28
2.3	Detection of simulated sources as a function of two source parameters.	29
2.4	Comparison of the detected fraction of sources between this work and S07.	31
2.5	Comparison of detected sources between this work and S07 as a function of two source parameters.	32
2.6	Detected vs. input source parameters for this work and S07.	34
2.7	Fraction of sources detected after visual inspection of the data. . .	36
2.8	Comparison of measured to input source properties.	38
2.9	Comparison of measured to input axial ratios.	40
2.10	Comparison of measured to input sources position.	41
2.11	Distribution of recessional velocities for HVCs.	43
2.12	Comparison of simulated sources to WvW isolation criterion. . . .	44
2.13	Comparison of simulated sources to ALFALFA isolation criteria. .	46
3.1	Spatial distribution of UCHVCs and α .40 HVCs.	49
3.2	The measured properties of the UCHVCs detected only by the algorithm of this work compared to the full sample of UCHVCs. . .	52
3.3	The relative isolation of LG galaxies and α .40 HVCs from the large HVC complexes of WvW.	55
3.4	The isolation of LG galaxies and UCHVCs from the general α .40 HVC population.	58
3.5	Defining the most-isolated subsample.	60
3.6	Maps of the HI column density of the UCHVCs derived from ALFALFA spectral grids.	62
3.6	Continued	63
3.6	Continued	64
3.7	Histograms of measured properties for the UCHVCs.	73
3.8	The distribution of inferred properties for the UCHVCs.	79
3.9	The distribution of UCHVCs relative to the Magellanic Stream from N10.	84
3.10	A zoomed in view of the fall UCHVCs relative to the MS from N10. .	85
3.11	The distribution of UCHVCs in $v_{LSR}-l$ space compared to the compact cloud populations of GALFA.	89
3.12	Comparison of HI properties for the UCHVCs detected and not visible in the GALFA-HI dataset.	92
3.13	The HI content of LG galaxies.	100

3.14	Cosine of the angle from the Solar apex versus heliocentric velocity.	102
3.15	The distribution of subhalos from the Via Lactea II simulation compared to the UCHVCs.	105
4.1	The ALFALFA HI contours and spectrum of AGC268069.	113
4.2	Bandpass response of all spectral windows as a function of frequency.	118
4.3	Noise for three different methods of combining the SPW coverage.	119
4.4	Beam size and RMS as a function of different robustness values. .	122
4.5	Spectra of AGC268069 for all three imaging schemes and velocity resolutions.	124
4.6	Spectra of AGC268069 for the 16kHz data cubes after removing the pedestaled flux level.	126
4.7	Moment zero maps for all three imaging schemes.	128
4.8	Single-dish spectra of AGC268069.	130
4.9	Comparison of VLA, ALFALFA and LBW spectra.	132
4.10	Velocity field of AGC268069	133
4.11	Overlay of ALFALFA and VLA HI contours on SDSS image.	135
4.12	Spectra of the three cores and northern extension.	137
5.1	The distribution of \tilde{N}_{HI} for the $\alpha.40$ UCHVCs and the targets selected based on high surface brightness for optical imaging.	143
5.2	The distribution of angular diameters for the $\alpha.40$ UCHVCs and those selected for compactness for optical imaging.	144
5.3	The distribution of v_{LSR} for the $\alpha.40$ UCHVCs and the targets selected for optical imaging based on large recessional velocities. . .	145
5.4	The distribution of v_{LSR} for all UCHVCs considered as optical targets and the targets selected for optical imaging based on large recessional velocities.	146
5.5	The CMD filter for the two extremes of the distance range considered in the search for a stellar population.	160
5.6	An illustration of the source detection methodology used on the field of AGC268069.	161
5.7	The 4' smoothed source distribution for all distance moduli considered.	163
5.8	Filtered sources for a distance modulus of 24.4 filtered to four different spatial scales.	164
B.1	Dirty continuum image before and after continuum source subtraction.	202

CHAPTER 1

INTRODUCTION

The concordance Λ CDM cosmology does an excellent job of matching observations of the Universe on large scales, including large scale structure of galaxies and fluctuations in the cosmic microwave background. However, on small scales there exist discrepancies between predictions based on Λ CDM and observations of low mass galaxies. These discrepancies can be studied in depth in the Local Group (LG).

The LG is a unique observational environment due to its proximity – intrinsically faint galaxies can be studied in great detail, including resolved stellar population studies. This has led to the development of “near field cosmology” – the study of the LG and its galaxies in great detail as a test of cosmological models on small scales.

The classical dwarf galaxies of the LG, discovered before the advent of wide-field optical surveys using CCDs, consist of 38 systems [Mateo, 1998]. The dwarf galaxies are either a satellite galaxy bound to the Milky Way (MW) or M31 or a general member of the LG isolated from the massive spiral galaxies. There is clear morphological segregation in the LG with the bound satellites being gas-poor dwarf spheroidals (dSphs) and dwarf ellipticals (dEs) while the isolated LG dwarfs are generally gas-rich dwarf irregulars (dIs) [Grebel and Gallagher, 2004].

1.1 The Small Scale Crisis

The first N-body simulations of the LG revealed a factor of ~ 100 more satellite galaxies than observed in the LG [Klypin et al., 1999, Moore et al., 1999]; this mismatch in numbers is often referred to as the “missing satellites problem” from the title of one of the early papers. This discrepancy was also universally predicted by semi-analytic galaxy formation models as a mismatch between the dark matter halo mass function and the galaxy luminosity function at the low mass end [Kauffmann et al., 1993]. Recent surveys in the optical and radio and continue to see this discrepancy as a difference between the predicted slope for the mass function of dark matter halos versus the observed slopes of the luminosity function [Blanton et al., 2005], the neutral hydrogen (HI) mass function [Martin et al., 2010], and the velocity width function [Papastergis et al., 2011].

In addition to the lack of observed low mass galaxies, there are outstanding challenges in matching the observed properties of low mass galaxies to dark matter halos from simulations. The rotation curves in dwarf and low surface brightness (LSB) galaxies do not rise as fast as predicted for dark matter halos [e.g. Moore, 1994, McGaugh et al., 2001, de Blok et al., 2001b, Blais-Ouellette et al., 2001, de Blok and Bosma, 2002, Marchesini et al., 2002]. When rotation curves are converted to density profiles of galaxies, the result is that dwarf and LSB galaxies show the presence of a $\sim \text{kpc}$ -sized constant density core while dark matter simulations predict that dark matter halos have a cuspy core [de Blok et al., 2001a, de Blok and Bosma, 2002]. de Blok et al. [2003] definitively examined the observational biases that could affect the determination of rotation curves for galaxies and find that the differences between the observed and predicted rotation curves

for dwarf galaxies are real.

The most recent mismatch between simulations and observations is the “too big to fail” problem. The naive expectation is that the most massive subhalos would host the luminous satellites of the MW. However, high resolution simulations of MW analogs find that the properties of the most massive subhalos are inconsistent with the luminous dSphs of the MW. For example, the most massive subhalos in simulations are too dense to match observations of dSphs [Boylan-Kolchin et al., 2011] and have larger maximum circular velocities (i.e., total mass) than the observed dSphs [Boylan-Kolchin et al., 2012]. Generally it appears that the most massive subhalos of simulations are more massive than the observed luminous satellite galaxies.

1.2 Solving the Problem

The discrepancies outlined in the previous section are collectively referred to as the “small scale crisis” in cosmology. In the case of the missing satellites problem, there are two clear ways to solve the problem: find the missing galaxies or motivate a modification to simulations that produces fewer low mass galaxies. Modifications to simulations should also address the other aspects of the small scale crisis, and this serves as a constraint for any new or modified physics. The two main ways to produce fewer luminous galaxies in simulations are to change the behavior of dark matter to be “warm”, avoiding clustering on small scales, or to incorporate baryonic physics and feedback to more accurately trace the evolution and location of baryons, with the net result of substantial baryon loss in the

lowest mass halos. Generally, the second approach to modifying simulations is accepted as the solution to the “small scale crisis”. There are (at least) three major effects to consider when accurately modeling baryons.

1.2.1 Relative Velocity of Dark Matter and Baryons

Recently, Tseliakhovich and Hirata [2010] pointed out that there is a relative velocity offset between baryons and dark matter early in the Universe and that this relative motion leads to a suppression of the first bound objects. Subsequent work has expanded on this to find that the bulk motion of the baryons suppresses the low mass halos that can host luminous galaxies [Fialkov et al., 2012, Bovy and Dvorkin, 2013], helping to bring predictions for low mass halos into agreement with observations of low mass galaxies. In addition, this bulk motion can delay reionization in some parts of the Universe as the low mass halos that are the sites of first star formation are not able to accrete baryons [Maio et al., 2011].

1.2.2 Reionization

After the formation of the first stars and galaxies, UV radiation floods the Universe, ionizing and heating the intergalactic medium (IGM) to a temperature of $\sim 10^4$ K. Reionization impacts small scale structure by preventing low mass halos from becoming luminous through photoevaporation of gas and by preventing the collapse of gas onto halos. Barkana and Loeb [2007] show that halos with circular velocities less than 10 km s^{-1} can lose all their gas to photoevaporation while halos with circular velocities up to 30 km s^{-1} can have substantial gas loss from

photoevaporation. In addition, heating the gas raises the Jeans mass and halos with circular velocities less than 30 km s^{-1} can no longer accrete fresh gas after reionization. Hoesft and Gottlöber [2010] find a characteristic mass of $6 \times 10^9 h^{-1} M_{\odot}$ for the transition between halos that are not affected by reionization and those with substantial baryon loss (but their simulations show that these halos may still contain some observable baryons).

1.2.3 Star Formation Feedback

Star formation feedback, including supernova (SN) explosion and winds from star formation, can play an important role in determining the baryon content of low mass galaxies. Ferrara and Tolstoy [2000] find that dwarfs with total masses $\lesssim 5 \times 10^6 M_{\odot}$ have all their gas blown away while dwarfs with masses up to $\simeq 10^9 M_{\odot}$ lose mass in outflows. In addition to resulting in depleted baryon content (and contributing metals to the intergalactic medium), SN feedback plays an important role in determining the mass profile of dwarf galaxies.

Navarro et al. [1996] and Mashchenko et al. [2008] initially showed that SN-driven winds expel large amounts of gas from low mass galaxies, resulting in a reduced central density, helping to explain the observed density profiles. More recently, Zolotov et al. [2012], Governato et al. [2012] and Teyssier et al. [2013] have shown that a *bursty* star formation history (SFH) can create substantial cores in low mass dark matter halos. However, these simulations do find that extremely low mass halos ($M_{vir} < 10^9$ or $M_{star} < 10^{4.5} M_{\odot}$) do *not* have a cored profile as they cannot host the bursty SFH necessary to drive the density profile change.

A cored density profile affects the evolution of low mass halos as they become more susceptible to tidal stripping. Brooks et al. [2013] modified the density profiles of subhalos in simulations based on the results of the aforementioned work. This correction enhances tidal stripping of the subhalos, and they find agreement in the total number of luminous satellite galaxies with observations of the dSphs of the MW. However, Peñarrubia et al. [2010] find that the dSphs have interacted substantially with the MW in the past, and that they need a cuspy profile to have avoided substantial tidal stripping, and Boylan-Kolchin et al. [2012] find that the masses of the dSphs are too low for the SN-driven processes to have a substantial effect on the density profile. Simulating the processes that drive and affect baryon evolution is an active area of work, aided by observations that can place constraints on the simulations.

1.3 Constraining the Problem by Finding New Galaxies

One way to resolve the missing satellites problem is to find some of the missing galaxies. Newly discovered systems missed by previous observations will presumably inhabit a parameter space of extremely low luminosities, surface brightnesses and masses. These systems provide laboratories for testing and constraining simulations as they are strongly affected by baryon loss and redistribution.

1.3.1 Ultra-faint Dwarf Galaxies

The census of classical LG dwarf galaxies was limited by sensitivity to low surface brightness galaxies. The advent of wide-field optical surveys, specifically the Sloan Digital Sky Survey (SDSS), increased the number of known MW satellites with the discovery of the ultra-faint dwarf galaxies (UFDs). The UFDs have luminosities from $10^2 - 10^5 L_{\odot}$, half-light radii from 20-350 pc and M/L ratios of 100 to over 1000, total masses within the baryon extent of $10^6 - 10^7 M_{\odot}$, generally old stellar populations, and are located at distances of tens to a few hundred kpc from the MW [Martin et al., 2008, Simon and Geha, 2007]. The name ultra-faint is well earned – the total luminosities of these objects are comparable to those of globular clusters, but they are clearly galaxies as their kinematics indicate they are dark matter dominated [Simon and Geha, 2007].

The discovery of UFDs is exciting and opens many possibilities into addressing the fundamental questions of how marginal galaxies form. Brown et al. [2012] present deep color magnitude diagrams (CMDs) from the Hubble Space Telescope (HST) for three UFDs. Their preliminary analysis of the SFH of these systems indicates that there is no intermediate-aged population and the stellar ages are within ~ 1 Gyr for all three systems, indicating that a global event, such as reionization, terminated the star formation.

Unfortunately, the vast majority of UFDs are located within the virial radius of the MW, making it difficult to understand which of their properties are intrinsic and which are the result of interaction with the MW. Bovill and Ricotti [2011] predict based on simulations that the vast majority of UFDs have been modified by tides; this is supported by observational evidence of tidal disruption [Simon and

Geha, 2007, Muñoz et al., 2010, Sand et al., 2012]. This makes it nearly impossible to determine which of the UFD properties, such as sizes and kinematics, are the result of evolution and which are result of environmental influence from interaction with the MW. Bovill and Ricotti [2011] predict the existence of ~ 100 fossil galaxies with luminosities less than $10^6 L_{\odot}$ that have remained isolated from the MW at distances of 400 kpc to 1 Mpc. A fossil galaxy is a system that formed a stellar population before reionization and has not merged with other galaxies since then; it is a relic of the first galaxies that formed. The properties of these systems would be the result of isolated evolution and could be used to constrain the astrophysical processes that impact baryon content in low mass halos.

1.3.2 Leo T: A Gas-Rich Ultra-faint Dwarf

One UFD discovered in SDSS is of particular note as it may represent a fossil galaxy. Leo T lies at distance of 420 kpc, safely outside the virial radius of the MW and was, until recently, the only gas-rich UFD discovered. Leo T is a star-forming galaxy with a HI mass of $2.8 \times 10^5 M_{\odot}$, an HI diameter of 600 pc, an indicative dynamical mass within the HI extent of $\sim 3.3 \times 10^6 M_{\odot}$, a total-mass-to-light ratio within the HI extent of 56, and a stellar mass of $\sim 1.2 \times 10^5 M_{\odot}$ [Ryan-Weber et al., 2008]. Given its gas content and distance, Leo T likely represents an unperturbed UFD, allowing environmental effects to be disentangled from the evolution of the lowest-mass galaxies. Indeed, Rocha et al. [2012] argue that Leo T is on its first infall to the MW. Leo T is on the edge of detectability for SDSS; were it located further away, its stellar population would not have been detected [Kravtsov, 2010]. UFDs with properties similar to Leo T but located further from the MW or with

fainter stellar populations would have been overlooked in the automated searches of SDSS. However, the HI content of Leo T would be detectable in a sensitive, wide area HI survey, raising the possibility that more isolated, gas-rich UFDs await discovery.

1.4 High Velocity Clouds as Dark Galaxies

The idea that galaxies can exist with HI as the only easily detectable component is not a new one; high velocity clouds (HVCs) have been considered many times as possible tracers of dark matter halos. HVCs are associations of neutral hydrogen with anomalous velocities compared to those expected for a differentially rotating Galactic disk. Since their discovery by Muller et al. [1963], much thought has been given to understanding the origin of the HVCs. Oort [1966] was among the first to critically consider different origins, and Wakker and van Woerden [1997] offer a thorough review. Possible origins include cold gas accretion, a Galactic fountain model and gas in dark matter halos. As HI phenomena, HVCs range from relatively isolated clumps less than a degree to vast complexes spanning hundreds of degrees; it is clear that they likely arise from a variety of origins. Determining the nature of HVCs is hampered by the lack of direct distance information.

One intriguing hypothesis has been that HVCs represent dark matter dominated “dark” galaxies. Considered by Oort [1966], this scenario was originally ruled out as it could not explain a majority of the observed HVCs at the time. In their review, Wakker and van Woerden [1997] summarize the objections against the galaxy hypothesis for HVCs. The HVCs considered at the time were rela-

tively bright and hence would correspond to relatively massive objects if at LG distances. As massive objects, the lack of detection of similar HI clouds in other galaxy groups and narrow linewidths (low dynamical masses) are difficult to explain. In addition, the velocity distribution of the HVCs considered was offset from the velocity distribution of the LG galaxies, pointing to a different dynamical origin.

With the advent of large-scale, sensitive, blind HI surveys, interest was revived in HVCs as tracers of dark matter halos. Blitz et al. [1999] and Braun and Burton [1999] both postulated a LG origin for HVCs; Braun and Burton [1999] specifically proposed that compact HVCs (CHVCs), identified by their isolation and undisturbed spatial structure, were good candidates to represent dark matter halos throughout the LG. Further work since the discovery of the CHVC population suggests that they most likely represent a circumgalactic population; the HI masses and diameters of these objects are larger than theoretical and observational predictions. Section 3.4.1 contains an in-depth discussion on the nature of CHVCs.

1.5 Ultra-compact High Velocity Clouds as Gas-Rich UFDs

The Arecibo Legacy Fast ALFA (ALFALFA) HI line survey is the first blind HI survey capable of robustly surveying the LG for gas-rich low mass dark matter halos. Exploiting the huge collecting area of the Arecibo 305m telescope¹, ALFALFA has the sensitivity to detect $10^5 M_{\odot}$ of HI with a linewidth of 20 km s^{-1} at

¹The Arecibo Observatory is operated by SRI International under a cooperative agreement with the National Science Foundation (AST-1100968), and in alliance with Ana G. Méndez-Universidad Metropolitana, and the Universities Space Research Association.

1 Mpc. This sensitivity limit represents the ability to detect objects with HI and dynamical masses comparable to Leo T throughout the volume of the LG.

1.5.1 Ultra-compact High Velocity Clouds in ALFALFA

The increased sensitivity and resolution of ALFALFA has resulted in the detection of ultra-compact high velocity clouds (UCHVCs). Giovanelli et al. [2010, hereafter G10] originally discussed a set of UCHVCs that were consistent with being gas-bearing low mass dark matter halos at ~ 1 Mpc; this interpretation of the UCHVCs was referred to as the minihalo hypothesis. In this thesis, I expand on this work and present a catalog of UCHVCs for the current 40% ALFALFA data release, termed $\alpha.40$ [Haynes et al., 2011]. These objects have HI diameters of $4' - 20'$, HI integrated flux densities of $0.1 \text{ Jy km s}^{-1} - 8 \text{ Jy km s}^{-1}$, and velocity widths of $15 - 70 \text{ km s}^{-1}$. If located at distances of ~ 1 Mpc, they have HI masses of $\sim 10^5 - 10^6 M_{\odot}$, diameters of $\sim 2 - 3 \text{ kpc}$ and indicative dynamical masses of $\sim 10^7 - 10^8 M_{\odot}$. I offer further details on the minihalo hypothesis for this class of objects, drawing special attention to the properties of Leo T and Leo P (see below). The small angular diameters and low integrated flux densities allow the UCHVCs presented here to overcome objections raised against previous HVC samples proposed to represent gas-rich galaxies in the LG.

1.5.2 Proof of Concept: Leo P

The recent discovery of the star forming dwarf galaxy Leo P from ALFALFA survey data shows that galaxies similar to Leo T in the Local Volume may be identi-

fied via their 21cm line emission [Giovanelli et al., 2013, Rhode et al., 2013, Skillman et al., 2013]. Leo P was discovered during the normal course of identifying HI detections within the ALFALFA survey when it was noticed that one UCHVC could be associated with an irregular, lumpy light distribution in the SDSS images [Giovanelli et al., 2013]. Follow-up optical observations resolved a stellar population and a single HII region, confirming that the UCHVC is in fact a low mass galaxy, Leo P. It is important to note that Leo P was confirmed to be a galaxy because its young, blue stellar population was barely visible in the SDSS images; without recent star formation, the underlying older population of Leo P would not have been visible at all in the SDSS images. Leo P was discovered by its HI signature, and its existence strongly argues that other very low mass and (nearly) starless objects are included among the ALFALFA UCHVCs.

1.6 Overview of the Thesis

This dissertation takes advantage of the sky coverage, increased sensitivity, and angular and spectral resolution of the ALFALFA HI survey to search for candidate optically faint gas-rich low mass galaxies in the nearby Universe that have been missed by extant optical surveys.

In Chapter 2, I present the signal extraction algorithm used to identify the UCHVCs and discuss the completeness of the resulting source detections. Chapter 3 presents a catalog of UCHVCs from the $\alpha.40$ footprint and discusses these objects as a class in the context of the minihalo hypothesis.

The focus then turns to further exploration of the minihalo hypothesis for the

UCHVCs via detailed follow-up observations. In Chapter 4 I present the results of high resolution HI synthesis imaging with the *Karl G. Jansky* Very Large Array (VLA). In Chapter 5 I motivate the selection of a subset of UCHVCs for follow-up optical imaging and present preliminary results for one system.

I summarize my findings and suggest future directions for understanding the UCHVCs in the Conclusion (Chapter 6).

The Appendices include the signal extraction code discussed in Chapter 2 and full details on the data calibration and reduction for the VLA observations presented in Chapter 4 and the WIYN/pODI observations presented in Chapter 5.

CHAPTER 2

AUTOMATED SIGNAL IDENTIFICATION OF ULTRA-COMPACT HIGH VELOCITY CLOUDS IN THE ALFALFA HI SURVEY

The first step to building a catalog and examining the minihalo hypothesis for the UCHVCs is to identify the UCHVCs within the ALFALFA dataset. The UCHVCs are distinguished from typical extragalactic HI detections by narrow linewidths and extended spatial sizes, motivating the development of a specialized signal extraction algorithm optimized for UCHVC identification. In this Chapter, I present a 3D matched filtering algorithm designed and implemented specifically for UCHVC identification. The standard ALFALFA source identification scheme is described, and the necessity of this algorithm is motivated. A brief summary of the methodology and implementation of the algorithm is given. The effectiveness of the algorithm is tested with the insertion of simulated sources. In addition, the simulated sources serve as a test of the accuracy of measured source parameters reported in the catalog in Chapter 3 and the impacts of the isolation criteria on the selection of UCHVCs.

2.1 ALFALFA, Source Identification, and UCHVCs

Before designing a specialized scheme for the UCHVCs, it is important to understand the ALFALFA dataset, the standard source identification scheme and what motivates the need for a separate procedure for the UCHVCs.

2.1.1 The ALFALFA HI Data

ALFALFA surveys the sky using a seven-feed multi-beam receiver (ALFA) in “drift” mode: the telescope is normally parked along the local meridian and 7 tracks (2 polarizations each) of spectral data of 4096 channels are acquired continuously and recorded at a 1 Hz rate as the sky drifts by. All regions of the sky are visited twice with the two visits typically a few months apart in time. Upon completion of data taking of a region of the sky, data cubes of $2^\circ.4 \times 2^\circ.4$ in spatial coordinates are produced and sampled over a regular grid of $1'$ spacing in R.A. and Dec. After Hanning smoothing to 11 km s^{-1} resolution, the rms noise per channel of the data is typically 2 to 2.5 mJy per beam. The data cubes are stored in IDL structures referred to as grids; generally speaking, the data is separated into four different grids based on velocity range, and the lowest velocity grids (of interest for the UCHVCs) cover the velocity range $[-2000, 3000] \text{ km s}^{-1}$. The main component of the grid structure is the data array which contains a flux value (in mJy) for each polarization at every sampled (R.A., Dec., velocity) point. The spacing between adjacent grids (in both R.A. and Dec.) is 2° so that each grid overlaps with its contiguous neighbors by 0.2° ($12'$).

2.1.2 Source Identification in the ALFALFA Dataset

Saintonge [2007, hereafter S07] developed an automated signal extractor for the ALFALFA dataset. The algorithm of S07 is a one-dimensional matched filtering scheme; the philosophy of matched filtering is discussed in Section 2.2.1. The spectrum in each pixel of an ALFALFA grid is matched to a series of Hermite

polynomial templates. The detection of a galaxy requires the detection of spectra of similar velocity widths with a high significance in 5 or more contiguous pixels. After the generation of a preliminary source catalog from the automated algorithm, each source in the catalog is visually inspected and remeasured. The measurement tool fits ellipses to contours of constant flux density level and delivers a source position, given by the center of the ellipse encircling half of the total flux density of the source, source sizes (as the major and minor axes of said ellipse), flux density, velocity and linewidth.

2.1.3 Motivation for a Specialized Algorithm

In comparison with the extragalactic sources identified in the $\alpha.40$ catalog, the UCHVCs are typically spatially extended and have narrow velocity widths. Figure 2.1 illustrates this by showing the distribution of HI angular diameters (θ_{HI}) and velocity full width half maximums (FWHM; W_{50}) for the $\alpha.40$ extragalactic sources, $\alpha.40$ HVCs and the UCHVCs of this work. The UCHVCs are spatially extended compared to the extragalactic sources but generally small compared to the full HVC population of the $\alpha.40$ survey. The minimum velocity width used in the templates of the S07 identification algorithm is 30 km s^{-1} , the typical *maximum* width of the UCHVCs. For these reasons, a special source identification algorithm was developed for the UCHVCs in addition to the standard ALFALFA pipeline.

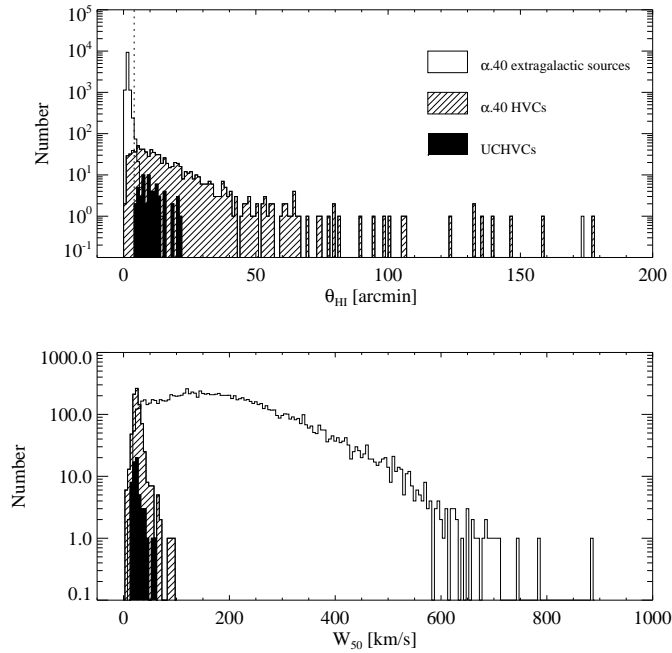


Figure 2.1: The distribution of HI angular diameters and full width half maximum of the HI line (W_{50}) for the UCHVCs (filled histograms), $\alpha.40$ sources classified as reliable extragalactic detections (unfilled histograms), and $\alpha.40$ HVCs (hashed histograms). The UCHVCs and HVCs occupy a small range of narrow velocity widths while the majority of extragalactic detections have broad lines.. The UCHVCs are spatially large compared to the extragalactic detections but generally small compared to the $\alpha.40$ HVCs.

2.2 Methodology of the Algorithm

The UCHVC signal extractor is based on the philosophy of S07 but optimized for UCHVCs with three main differences: a limited velocity range, three-dimensional matched filtering, and the use of Gaussian templates. Only a limited velocity range of the ALFALFA data set, $-500 < v_{\odot} < 1000 \text{ km s}^{-1}$, is selected as this is the expected velocity range for objects within the Local Volume. Because only a limited velocity range is examined, it is reasonable to perform a full three-dimensional matched filtering, matching both the spectrum of the source and the spatial position and size simultaneously. Gaussian templates are used to describe both the spatial extent and the velocity profile of the UCHVCs. The velocity profiles of the UCHVCs arise mostly from thermal support and are observed to generally be Gaussian in shape. The spatial profile of the UCHVCs is unknown, but a Gaussian captures the important behavior: an exponential decrease in the profile and a well-defined half-light radius.

2.2.1 Matched Filtering

In a matched filtering algorithm, the form of the signal is assumed. The unknown signal is then convolved with a series of templates; the highest convolution values corresponds to the best-matched template.

We are assuming sources that are 3D Gaussians, described by a velocity distribution of σ_{vel} and spatially symmetric with a size described by σ_{sp} . Our assumed

signal has the form

$$s(x, y, z) = A_s \exp(-(x - x_0)^2/2\sigma_{s,sp}^2) \exp(-(y - y_0)^2/2\sigma_{s,sp}^2) \exp(-(z - z_0)^2/2\sigma_{s,vel}^2), \quad (2.1)$$

where (x_0, y_0, z_0) describe the location of the signal within a data cube and $\sigma_{s,sp}$ and $\sigma_{s,vel}$ represent the spatial and velocity extent of the signal, respectively. The templates have a similar form with a normalization of $1/\sqrt{\pi^{1/2}\sigma_t}$ for each Gaussian component. Explicitly, the templates have the form of

$$t(x, y, z) = \frac{1}{\sqrt{\pi^{3/2}\sigma_{t,sp}^2\sigma_{t,vel}}} \exp(-x^2/2\sigma_{t,sp}^2) \exp(-y^2/2\sigma_{t,sp}^2) \exp(-z^2/2\sigma_{t,vel}^2) \quad (2.2)$$

for various values of $\sigma_{t,sp}$ and $\sigma_{t,vel}$. We wish to convolve the templates with the unknown signal to find the best-matched template. As a convolution corresponds to a multiplication in the Fourier domain (via the convolution theorem), we first look at the Fourier transforms (FT) of the signal and templates.

$$\begin{aligned} S(k, l, m) &= A_s \sqrt{(2\pi)^3 \sigma_{s,sp}^4 \sigma_{s,vel}^2} \exp(-\pi^2 k^2 (2\sigma_{s,sp}^2)) \times \\ &\quad \exp(-\pi^2 l^2 (2\sigma_{s,sp}^2)) \exp(-\pi^2 m^2 (2\sigma_{s,vel}^2)) \times \\ &\quad e^{-2\pi i k x_0} e^{-2\pi i k y_0} e^{-2\pi i k z_0} \end{aligned} \quad (2.3)$$

$$\begin{aligned} T(k, l, m) &= \sqrt{(2/\sqrt{\pi})^3 \sigma_{t,sp}^2 \sigma_{t,vel}^2} \exp(-\pi^2 k^2 (2\sigma_{t,sp}^2)) \times \\ &\quad \exp(-\pi^2 l^2 (2\sigma_{t,sp}^2)) \exp(-\pi^2 m^2 (2\sigma_{t,vel}^2)) \end{aligned} \quad (2.4)$$

Then the product of the FTs of the signal and template are:

$$\begin{aligned} S \times T &= A_s \sqrt{(2^2 \sqrt{\pi})^3 \sigma_{s,sp}^4 \sigma_{t,sp}^2 \sigma_{s,vel}^2 \sigma_{t,vel}^2} \exp(-\pi^2 k^2 (2\sigma_{s,sp}^2 + 2\sigma_{t,sp}^2)) \times \\ &\quad \exp(-\pi^2 l^2 (2\sigma_{s,sp}^2 + 2\sigma_{t,sp}^2)) \exp(-\pi^2 m^2 (2\sigma_{s,vel}^2 + 2\sigma_{t,vel}^2)) \times \\ &\quad e^{-2\pi i k x_0} e^{-2\pi i k y_0} e^{-2\pi i k z_0} \end{aligned} \quad (2.5)$$

And the inverse FT gives the convolution of the signal and template:

$$\begin{aligned}
s \otimes t = & A_s \sqrt{\frac{2\sigma_{s,sp}^2 \sigma_{t,sp}}{\sigma_{s,sp}^2 + \sigma_{t,sp}^2}} \sqrt{\frac{2\sigma_{s,sp}^2 \sigma_{t,sp}}{\sigma_{s,sp}^2 + \sigma_{t,sp}^2}} \sqrt{\frac{2\sigma_{s,vel}^2 \sigma_{t,vel}}{\sigma_{s,vel}^2 + \sigma_{t,vel}^2}} \times \\
& \exp[-(x - x_0)^2 / 2(\sigma_{s,sp}^2 + \sigma_{t,sp}^2)] \exp[-(y - y_0)^2 / 2(\sigma_{s,sp}^2 + \sigma_{t,sp}^2)] \times \\
& \exp[-(z - z_0)^2 / 2(\sigma_{s,vel}^2 + \sigma_{t,vel}^2)].
\end{aligned} \tag{2.6}$$

Clearly, for any given template the maximum value of the convolution of the template with the signal occurs at (x_0, y_0, z_0) - the location of the signal. Then the question is, for what template is the convolution value the greatest? This can be checked by looking for extremum values of

$$A_{templ} = \sqrt{\frac{2\sigma_{s,sp}^2 \sigma_{t,sp}}{\sigma_{s,sp}^2 + \sigma_{t,sp}^2}} \sqrt{\frac{2\sigma_{s,sp}^2 \sigma_{t,sp}}{\sigma_{s,sp}^2 + \sigma_{t,sp}^2}} \sqrt{\frac{2\sigma_{s,vel}^2 \sigma_{t,vel}}{\sigma_{s,vel}^2 + \sigma_{t,vel}^2}} \tag{2.7}$$

with respect to both $\sigma_{t,sp}$ and $\sigma_{t,vel}$.

Looking first to $\sigma_{t,sp}$, we have:

$$\frac{\partial A_{templ}}{\partial \sigma_{t,sp}} = \frac{2\sigma_{s,sp}^2}{\sigma_{s,sp}^2 + \sigma_{t,sp}^2} - \frac{4\sigma_{s,sp}^2 \sigma_{t,sp}^2}{(\sigma_{s,sp}^2 + \sigma_{t,sp}^2)^2} \tag{2.8}$$

By setting the above equation equal to zero and solving for $\sigma_{t,sp}$, we find $\sigma_{t,sp} = \sigma_{s,sp}$ for the maximum convolution value. Similarly for $\sigma_{t,vel}$, we have:

$$\frac{\partial A_{templ}}{\partial \sigma_{t,vel}} = \frac{1}{2} \sqrt{\frac{\sigma_{s,vel}^2 + \sigma_{t,vel}^2}{2\sigma_{s,vel}^2 \sigma_{t,vel}^2}} \left(\frac{2\sigma_{s,vel}^2}{\sigma_{s,vel}^2 + \sigma_{t,vel}^2} - \frac{4\sigma_{s,vel}^2 \sigma_{t,vel}^2}{(\sigma_{s,vel}^2 + \sigma_{t,vel}^2)^2} \right) \tag{2.9}$$

which results in $\sigma_{t,vel} = \sigma_{s,vel}$ for the maximum value.

Then, the maximum convolution value occurs when the template parameters exactly match those of the signal and has a value of $A_s \sigma_{s,sp} \sigma_{s,vel}^{1/2}$.

2.3 Implementation

This section details the implementation of the matched filtering algorithm for the ALFALFA HI data. First the chosen templates are justified. Then there is an explanation of how the presence of Galactic hydrogen and galaxies is handled. Finally, this section ends with a description of how sources are identified within the ALFALFA data using the automated algorithm and how the final source properties are arrived at through visual inspection and measurement.

2.3.1 The Templates

Gaussian templates are used to describe both the spatial extent and the velocity profile of the UCHVCs. The templates range from a spatial FWHM (θ_{HI}) of 4' to 12' in steps of 2' and velocity FWHM (W_{50}) of 10 km s⁻¹ to 40 km s⁻¹ in steps of 6 km s⁻¹. The lower bound of the spatial templates is set by the beam size of Arecibo. The upper size bound is near the median size value of the UCHVCs and represents our emphasis on detecting ultra-compact clouds. UCHVCs can be larger in size than 12' and the matched filtering of the 12' template to a UCHVC with HI diameter greater than 12' is robust (see Section 2.5). A W_{50} of 10 km s⁻¹ represents the narrowest source that can be spectroscopically resolved in the ALFALFA data. The warm neutral medium is thought to be the dominant phase of the ISM in minihalos [e.g. Sternberg et al., 2002]; for a reasonable range of temperatures (6000 – 10000 K) for the warm neutral medium in the UCHVCs, thermal broadening results in linewidths of $\sim 16 - 21$ km s⁻¹. Thus for a cloud of 40 km s⁻¹ linewidth, we would expect the large scale motion to be ~ 34 km s⁻¹

for the warmest clouds, after subtracting the thermal broadening contribution in quadrature. For a typical size of $10'$ at an indicative distance of 1 Mpc, the dynamical mass based on this unbroadened linewidth is $\sim 10^8 M_\odot$. This is a reasonable upper limit to the dynamical mass we may expect to be traced out for a more massive dark matter halo of $\lesssim 10^{10} M_\odot$, and matches the dynamical mass traced by the baryon extent of the presumably more massive galaxies in the Survey of HI in Extremely Low-mass Dwarfs (SHIELD) sample [Cannon et al., 2011].

2.3.2 Galactic Hydrogen

The Galactic hydrogen is extremely bright and contains many small substructures. ALFALFA was designed as an extragalactic HI survey and is not geared towards dealing with the vast amounts of structure and high dynamic ranges present in the Galactic HI. The simplified approach taken here is to identify and blank the Galactic hydrogen before running the matched filter routine as otherwise the signal extraction algorithm would become bogged down by the bright structure in the Galactic HI. Future work will involve using a methodology developed for the GALFA-HI survey (a complementary ALFA survey designed to study the Galactic HI) by Saul et al. [2012] to search the Galactic HI within the ALFALFA survey for UCHVCs.

The Galactic hydrogen is identified by totaling the absolute value of each velocity slice in a data cube. The median value for all slices is identified, and any slice with a value greater than $(\text{median} + 0.15 \times \text{median})$ is identified as being Galactic hydrogen. Each data point of this velocity slice is then replaced with the value 0.00. It should be noted that this methodology will occasionally identify slices of

the data cube with extremely large and bright galaxies or HVCs as being Galactic hydrogen. Generally, this is not of concern both because any clouds identified near bright galaxies or HVCs would be considered associated rather than isolated structures and because, as discussed below in Section 2.3.5, the data is visually inspected as source properties are measured by hand, allowing overlooked sources to be identified. This also allows bright isolated clouds within the Galactic hydrogen to be identified as potential minihalo candidates.

2.3.3 Detection of Galaxies

This algorithm does not distinguish between galaxies and the UCHVCs of interest. The templates used to identify (and then subtract) sources within the data cubes are designed to be well matched to the UCHVCs. Specifically, they have narrow velocity widths compared to what is typical for galaxies. As a result, galaxies typically show up as multiple detections as multiple templates are fit (and subtracted) across the width of the galaxy. While this impacts the run time of the code slightly, it does not affect the detection of UCHVCs and is typically a minor consideration as the volume searched for the UCHVCs is small and so few galaxies fall within it.

2.3.4 Identifying a Source

The methodology behind the signal extraction algorithm is described above in Section 2.2. The practical implementation of the code is that a $2^\circ.4 \times 2^\circ.4$ data cube is loaded. The cube is restricted to the $[-524, 1020]$ km s^{-1} range, and the Galactic

hydrogen is identified and masked to zero. Then a loop begins to do the matched filtering and identify the sources within the data. The loop is continued as long as the last identified source is above the S/N threshold set with the initial call to the algorithm; this threshold is typically set to 4.3. Briefly, the signal identification loop undertakes the following steps (the full code is located in Appendix A):

- The convolution of the data cube with each of the 25 templates is done via the fast Fourier transform (FFT)
- The location of a signal is identified as being the highest convolution value for all templates.
- The best matched source properties are determined by interpolating the convolution values of all the templates at the location of the signal using a two-dimensional four-degree polynomial.
- The peak and integrated flux of the source are derived from the data.
- It is checked that the source is strongly detected in at least 5 pixels.
- The S/N of the detected source is calculated.
- The difference between the two polarizations is checked.
- The source is subtracted from the data.
- If the source is above the S/N threshold and strongly detected in at least 5 pixels, it is added to the source catalog and the loop continues. If the source has a polarization difference or is in a low weights region, a flag is activated in the source catalog.
- If the source is below the S/N threshold, the algorithm writes the current source catalog to file and exits.

2.3.5 Measuring Source Properties

As in the standard ALFALFA pipeline, the automated signal identification algorithm is first run over each data cube, producing a preliminary source catalog. Then each source in the catalog is visually inspected and remeasured. The measurement tool returns a source position, source sizes, flux density, velocity and linewidth. During this stage, the data cubes are stepped through in velocity, and the Galactic hydrogen is examined for any clearly isolated small clouds.

2.4 Creation of Test Data

In order to test the performance of the signal extraction algorithm, 100 random grids were selected. These 100 grids all had 3 random artificial sources added to them. The details of how the artificial sources were simulated are below. After seeding the sources in the various grids, both the signal extraction algorithm described here and that presented in S07 were run over the grids. The found sources were recorded and automatically compared to the known simulated sources for the analysis in the following sections.

2.4.1 Simulating Artificial Sources

The location of the artificial source is randomly selected. The sources are restricted to the velocity range $[-524, 1020]$ km s⁻¹ and to lie within the central 2° of the data cube. The velocity restriction corresponds to the expected range of velocities for objects within the Local Volume and is the range of velocities considered in the

signal extraction algorithm. The sources are restricted to the central 2° of the data structure as sources that lie closer to the edge would be detected and measured in the contiguous grid where the source lies closer to the center of the grid.

After determining the location of the source, the source properties are randomly selected. Both the velocity and spatial extent are modeled as Gaussians. The W_{50} value is randomly selected from the range $[10, 40] \text{ km s}^{-1}$ and θ_{HI} is selected from the range $[4', 30']$, with up to a 5% difference in extent between the R.A. and Dec. directions. The W_{50} range corresponds to the range of templates used by the signal extractor; see Section 2.3.1 for a discussion of why this range of W_{50} values is considered. The range of θ_{HI} is set by the beam size of Arecibo on the lower end and by the maximum size of a cloud considered to be a UCHVC on the upper end. The amplitude of the source is randomly selected from the range $[0.5, 16] \text{ mJy}$ as testing indicates this a good range to have both strong sources and sources at the noise level.

After determining all the properties of the source, it is added to the data cube. The integrated and peak fluxes and the S/N of the source are then calculated from the data. These calculated values and the chosen properties of the source (location, W_{50} , θ_{HI}) are then recorded in an IDL structure for later comparative use.

2.5 The Completeness of the Algorithm

By comparing the properties of the detected sources to the non-detected sources, we can understand the limitations of the signal extraction algorithm and the completeness of the catalog of UCHVCs. Figure 2.2 shows the fraction of sources

detected (solid line) after running the signal extraction algorithm and the total number of sources (dashed line) as a function of various source properties. The S/N and recessional velocity of a source clearly have the strongest effect on the detection. One important result to note is that the fraction of sources detected is constant with spatial size; the maximum spatial size of the templates is $\theta_{HI} = 12'$ but the sources with θ_{HI} up to $30'$ are still robustly detected. The W_{50} , θ_{HI} , and cz plots show the uniform distribution of the simulated sources. The S/N plot shows many more low signal than high signal sources. While the amplitudes were drawn from a uniform distribution, the total signal of the source depends on the integrated flux density which depends on the size and velocity width of the source.

In Figure 2.3 the detections (solid circles) and non-detections (open squares) are plotted together as a function of two different source parameters. This allows us to examine how the ability to detect a source depends on related source parameters. The clearest trend is to see is that there is a dependency between W_{50} and S/N. Generally, sources with smaller W_{50} values are detected at lower S/N values than sources with larger W_{50} values. This is a well known effect in spectral line data, where the sensitivity depends upon the velocity width of the source. Narrower signals have more flux per channel and hence are easier to detect. There is a second effect also visible in this plot – there are more narrow width sources at low S/N than broader sources. This is a result of drawing the peak flux from a uniform distribution rather than the integrated flux density; the narrower sources have systematically lower integrated flux densities and hence lower S/N.

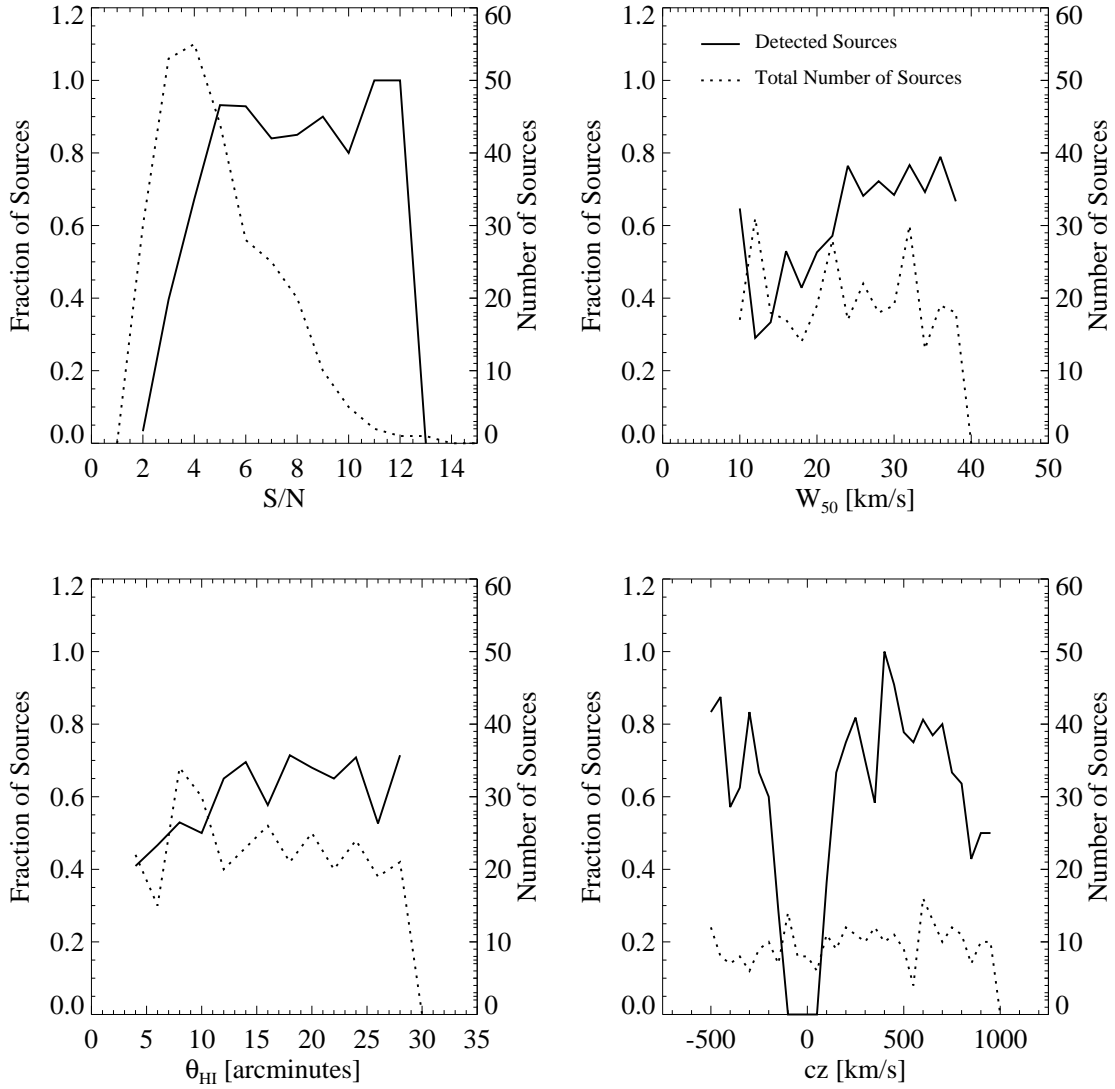


Figure 2.2: The solid lines show the fraction of simulated sources detected as a function of various simulated source properties; the dashed line indicates the total number of simulated sources. The fraction of sources detected falls off quickly for low signal sources and detection is relatively flat as function of W_{50} and θ_{HI} . Importantly, sources with $\theta_{HI} > 12'$ (the maximum template size) are robustly detected. The effects of disregarding the Galactic hydrogen are evident in the strong dip in detected sources at 0 km s^{-1} .

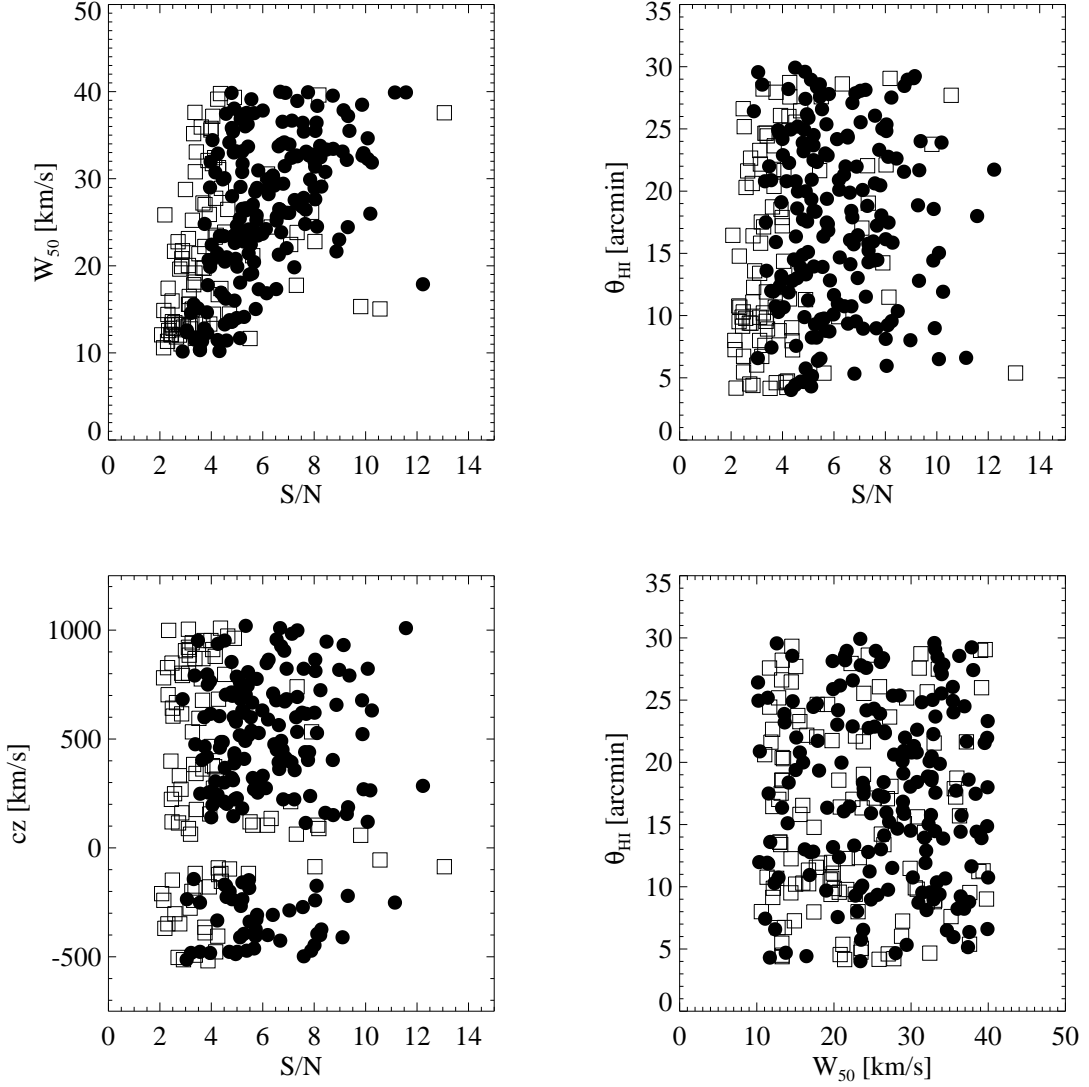


Figure 2.3: Plots comparing the detections (solid circles) to non-detections (open squares) as function of two different source parameters for simulated sources. The S/N of a source is the primary determinant in whether a source will be detected. In addition, the effect of linewidth on sensitivity is evident in the detection of more narrow linewidth sources with low signal.

2.6 Comparison to the Standard ALFALFA Scheme

There is a clear motivation in the expected properties of the UCHVCs for developing a specialized signal extraction algorithm; in this section we show directly that this algorithm detects sources missed by the standard ALFALFA data identification scheme. Figure 2.4 compares the fraction of sources detected by the algorithm presented here (solid line) and the fraction of sources detected by the algorithm of S07 (dashed line). Generally, the algorithm presented here outperforms that of S07 for all source parameter values. The exception is that the algorithm of S07 occasionally detects a source within the Galactic HI as result of the different handling of Galactic hydrogen emission. Figure 2.5 compares the detected sources for the two algorithms as a function of two source parameters. The algorithm of this work does a better job of detecting the extended sources, especially for narrow velocity widths and low S/N values. The algorithm of S07 can still detect some narrow velocity width sources as the matched filtering identification scheme is fairly robust. If there is enough signal, a source will be detected even if it is not matched perfectly to a template; this is how the algorithm presented here is robust to UCHVCs $> 12'$ in extent. However, in cases where source properties lie outside the range of templates used, the estimates of source parameters will not be accurate.

While visual inspection and interactive source measurement occurs after source identification by both algorithms, it is still useful to examine the reported sources properties of the two algorithms and see how they compare to the simulated source parameters. In Figure 2.6, the reported source properties of the algorithm of this work (crosses) and that of S07 (squares) are shown as a function

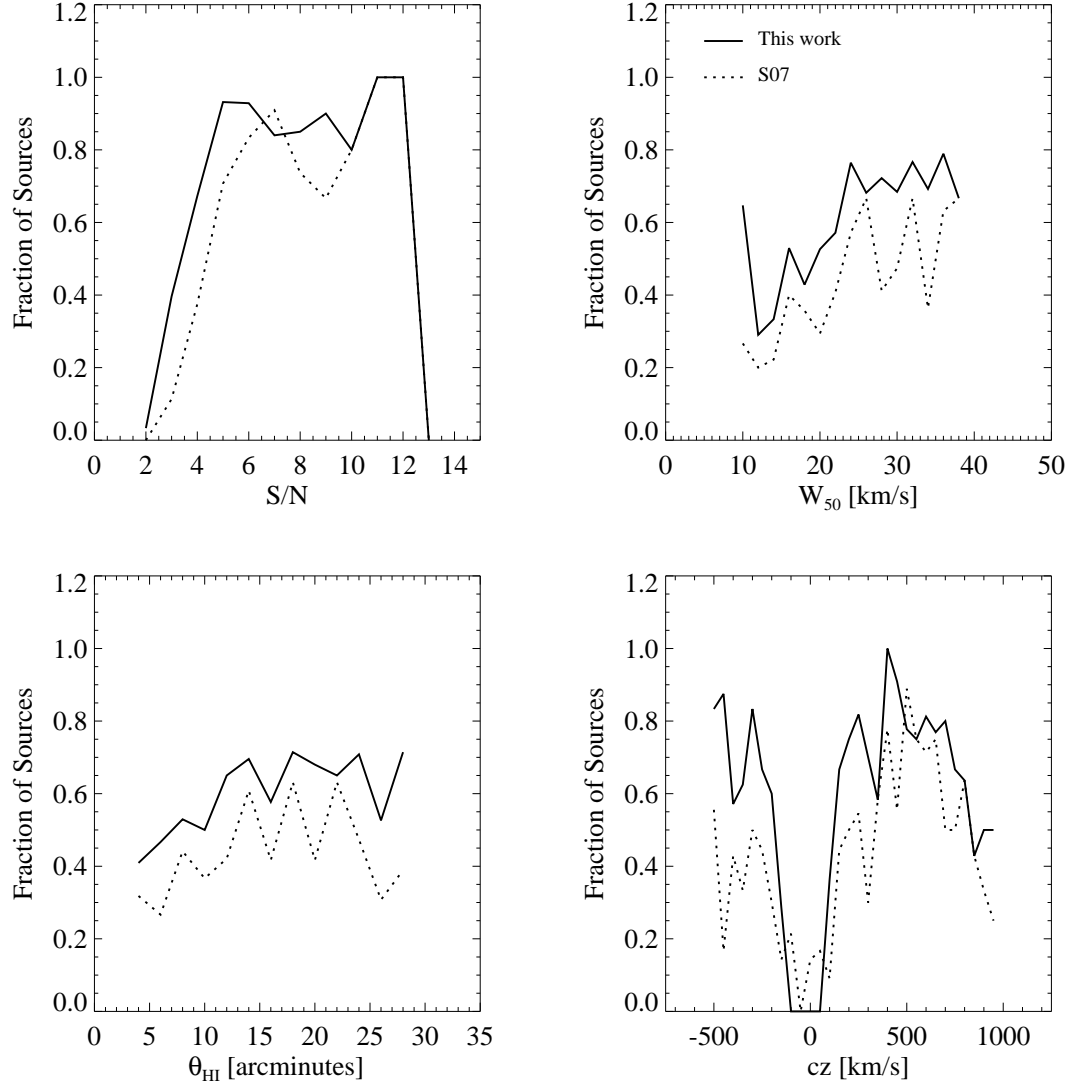


Figure 2.4: Fraction of simulated sources detected by this work (solid line) and that of S07 (dashed line) as a function of the simulated source properties. Generally this work outperforms that of S07, although S07 is able to detect a few sources within the Galactic HI.

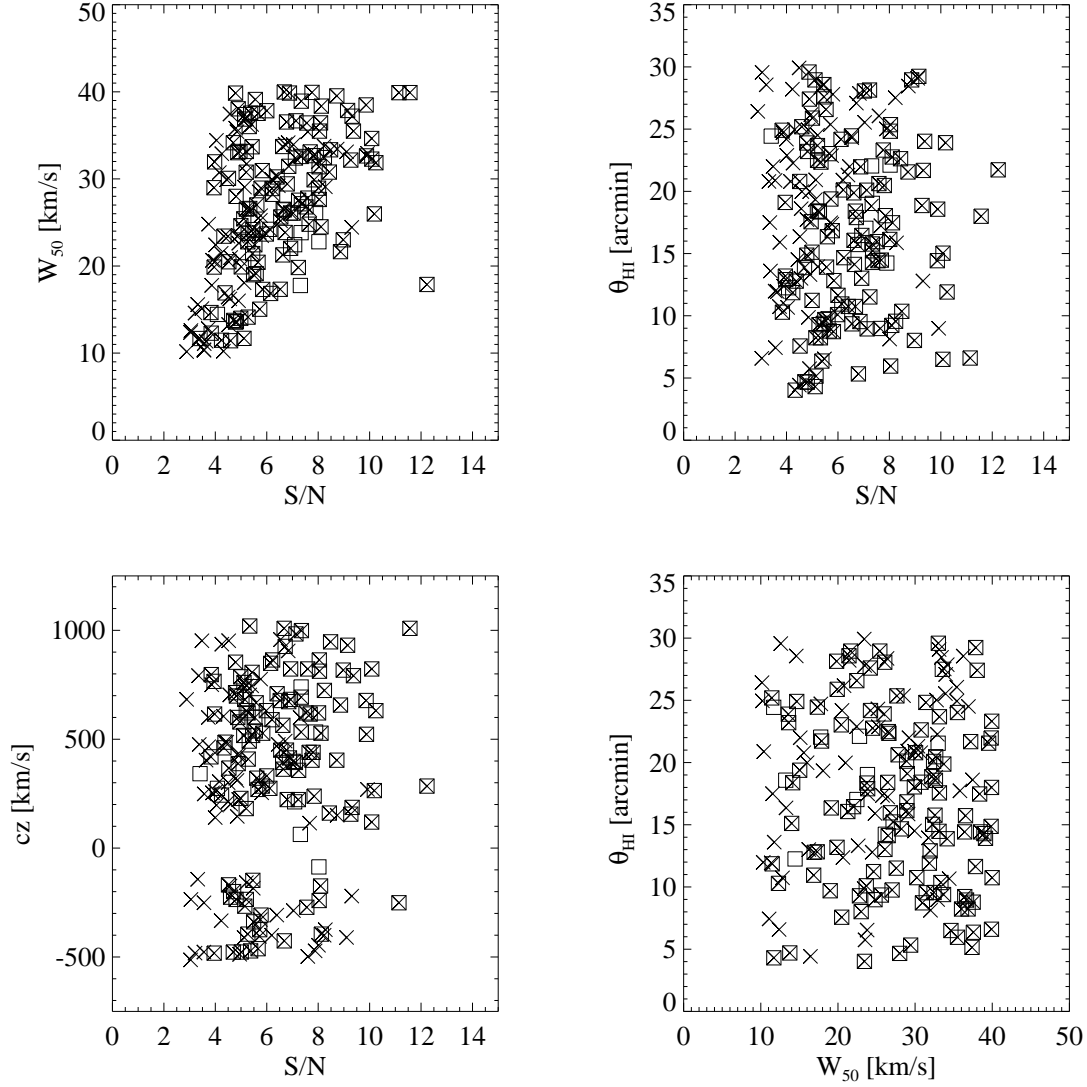


Figure 2.5: Detections of the algorithm presented here (crosses) and by the algorithm of S07 (open squares) as a function of two different input source parameters. This work is more sensitive to low S/N sources than S07 and is especially more sensitive to sources with low S/N that have $\theta_{HI} > 20'$.

of the simulated source properties. The algorithm of this work does a good job of determining the linewidths and spatial extents of the UCHVCs, up to the maximum spatial template considered. The algorithm of S07 is either hit or miss; it is able to assign sources sizes larger than $12'$ but often reports an extended source as a point source. While the algorithm of S07 is able to detect some narrow linewidth sources, the reported W_{50} is too large as it is limited to the smallest template used by S07. The algorithm presented here is able to outperform that of S07 for centroiding the sources.

2.7 Comparison to Visual Inspection

After running the signal extraction algorithm, data cubes are visually inspected and source properties are measured by hand. At this stage, sources missed by the signal extraction algorithm can be added to the catalog. I completed the detection scheme for the simulated sources by visually inspecting the data cubes and measuring their properties. Figure 2.7 shows the fraction of sources detected by visual inspection in comparison to the fraction detected solely by the automated algorithm. Visual inspection results in the slightly increased detection of lower S/N sources. Notably, the fraction of sources with W_{50} less than 20 km s^{-1} is substantially increased upon visual inspection; these sources often show strongly in only a single channel. The fraction of detected sources increases for all spatial sizes. There is a slight increase for the fraction of sources detected near the Galactic HI, but in general the presence of the Galactic HI still masks any sources that

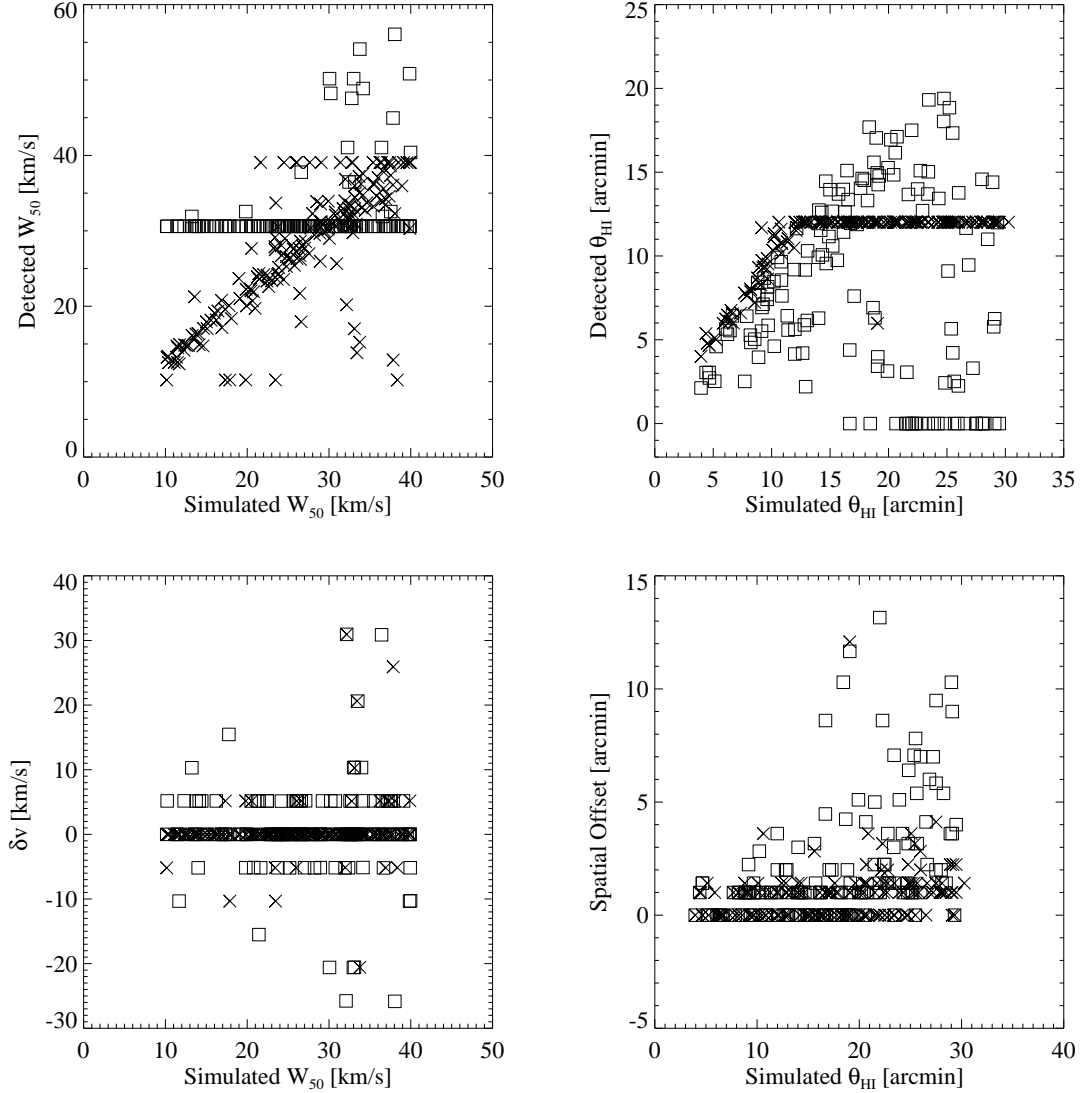


Figure 2.6: Detected versus simulated properties for sources for this work (crosses) and that of S07 (squares). Sources detected by S07 have a minimum W_{50} of 30 km s^{-1} as a result of the templates used, and sources detected by this work have a maximum θ_{HI} of $12'$ as that is the maximum spatial extent of the templates used. Within the range of templates used, the algorithm presented here does a good job of recovering the true source parameters.

lie at similar velocities.

It should be noted that there is one source which registered as an automated detection but is not considered a detection upon visual inspection; this source serendipitously lies at the same location as an actual UCHVC, and the automated algorithm detects the actual UCHVC; visual inspection of the data set does not reveal a noticeable second (simulated) source. The impacts of the isolation criteria (see Section 3.1.4) on the simulated sources are considered below in Section 2.9.

2.8 Accuracy of Reported Source Values

Since the simulated sources have been measured as part of testing the extraction algorithm, the measured properties can be compared to input properties in order to understand how well measured the source properties are. Figure 2.8 compares the measured source properties to the input simulated properties. The biggest difference is between the recorded S/N of the simulated source and the measured S/N; there is a monotonic relation between the measured and recorded S/N and the measured S/N is larger in all cases. This is a result of the differences in how the S/N of the source is calculated in the input data cube and how the measured S/N is computed. The measured S/N calculation is more sophisticated and uses the ellipse fitting to more accurately determine the peak flux of the source, as opposed to a rough estimate of the peak flux undertaken when simulating the source. This results in the measured S/N being systematically higher than the computed S/N. The other source properties lie on a one-to-one relation with some scatter. In the upper right corner, the comparison of measured to input W_{50} values

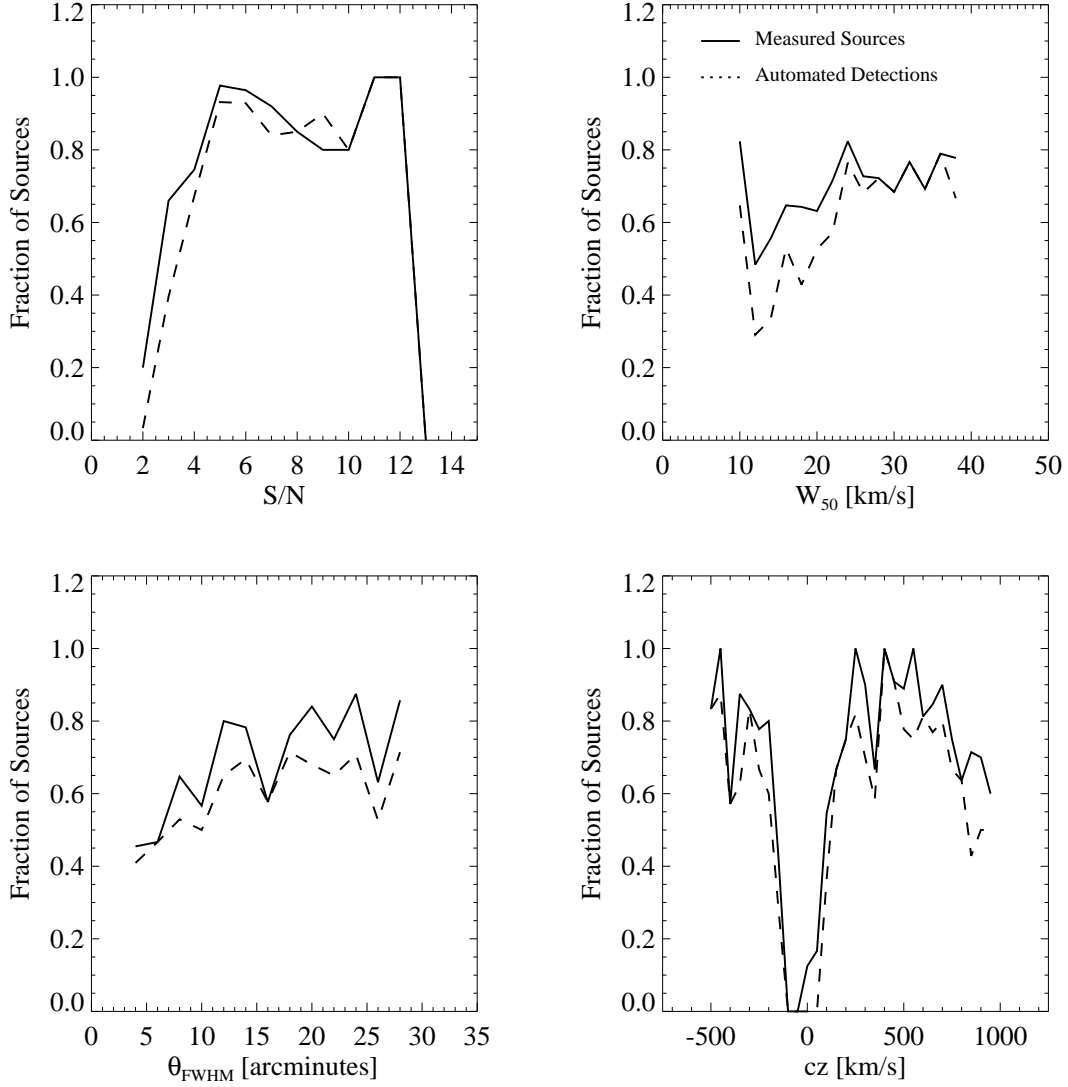


Figure 2.7: Fraction of sources detected after visual inspection (solid line) and those detected by the algorithm alone (dashed line) as a function of simulated source parameters. The inspection of the data reveals sources missed by the algorithm, including at Galactic velocities.

are shown along with the error bars recorded during measurement. Most sources have measured widths that match the input widths within the errors. Typically, the wider sources have larger error bars; these sources generally have lower S/N values and it is more difficult to determine where the edge of the line profile is. For θ_{HI} , the scatter between measured and simulated source properties increases with size.

The upper left panel of Figure 2.9 examines the difference between the input and measured axial ratios of the simulated sources. The simulated sources are near circular with axial ratios ranging from 0.95 – 1; this range of axial ratios is set by the 5% difference permitted in the extent of the source in the R.A. and Dec. directions. The measured axial ratios portray the sources as much more elliptical than they truly are with axial ratios distributed fairly evenly from 0.6–1 (and a few smaller axial ratios). The other three panels of Figure 2.9 show the dependence of the measured axial ratio on other measured source properties. The axial ratio shows no dependence on the size or width of the source but does show a slight dependence on the S/N of the source. The higher S/N sources typically have axial ratios closer to one (the true axial ratio value), and the range of measured axial ratios is much larger for lower S/N sources. Thus it appears that the highest S/N sources have measured axial ratios closest to the true values.

Given that the average measured θ_{HI} is well matched to the true input size, as can be seen in Figure 2.8, this tells us that the measured size of the sources is an accurate reflection of their true size but that the measured ellipticity values from the ALFALFA data should be treated carefully. A true understanding of source ellipticity will require follow-up observations. Since the measured axial ratio does depend upon S/N, deeper mapping observations with Arecibo could be used to

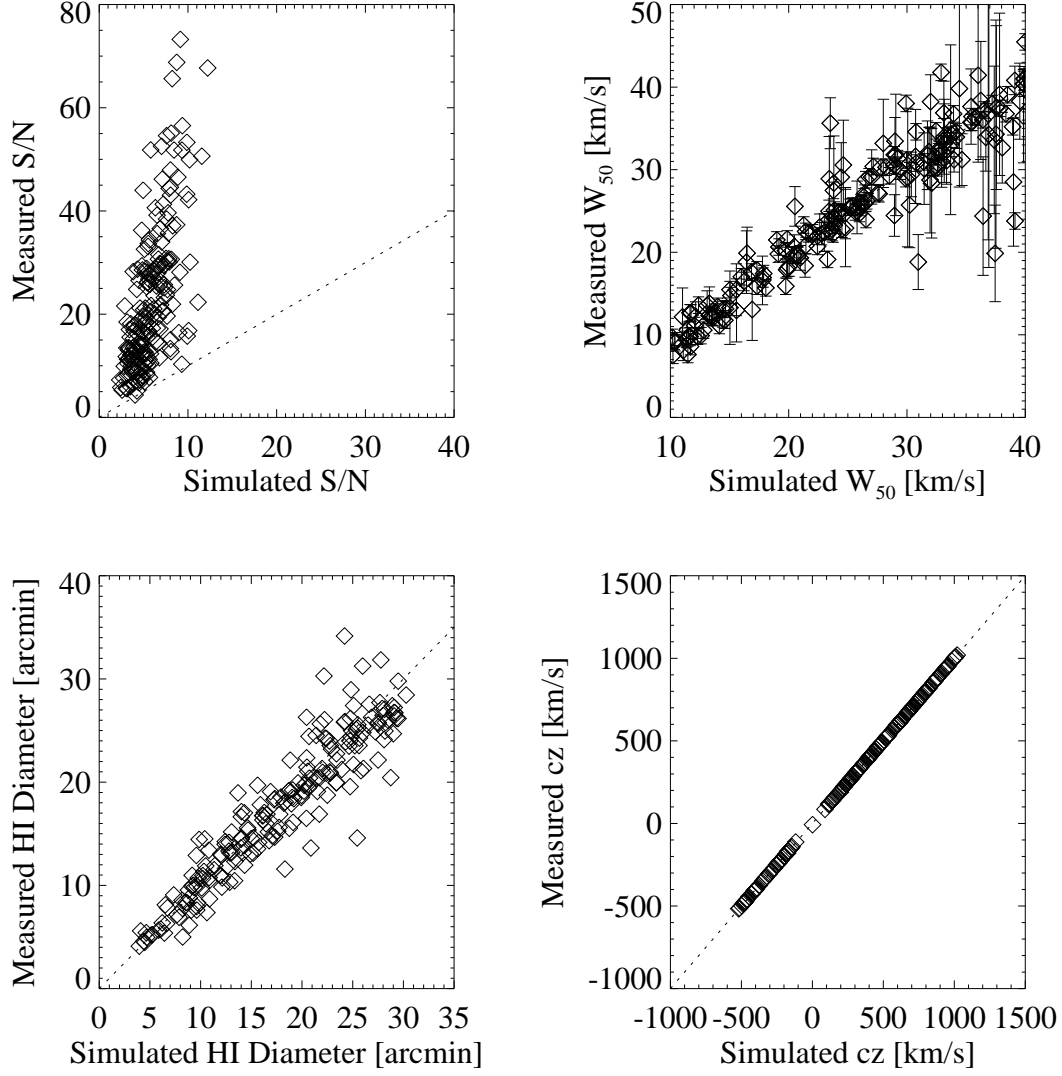


Figure 2.8: Comparison of measured to simulated source properties for S/N, W_{50} , θ_{FWHM} , and cz . Dotted lines indicate a one-to-one relationship in all cases. The large difference in S/N is a result of a more sophisticated calculation used by the measurement tool. The W_{50} , θ_{HI} , and cz values are in good agreement, indicating that the measured source parameters are an accurate representation of the true source properties.

place more accurate constraints on the ellipticity. Alternatively, higher resolution observations would also allow more accurate axial ratios to be measured.

Figure 2.10 examines the offset between the measured and expected position of the sources as a function of various source properties. The top row examines the spatial offset as a function of the size of the source and the S/N. Final positions of sources reported by the measurement tool include a pointing correction [Giovanelli et al., 2007, Kent et al., 2008]. The input position of the source does not have a pointing corrections applied, so some offset (on the order of $\lesssim 30''$) is naturally expected. Typically, the centroiding of the source is accurate to $\sim 1'$ with 70% of sources having this accuracy. Measured recessional velocities are extremely accurate with a maximum offset of 10 km s^{-1} and 80% of sources having accuracies within 2 km s^{-1} .

2.9 The Effects of the Isolation Criteria

As discussed in the presentation of the catalog of UCHVCs in Chapter 3, one of the defining criteria for a UCHVC is its isolation from other existing HVC structure. We discuss the importance of the isolation criteria and its definition in detail in Section 3.1.4. Here, we examine the simulated sources in the context of the isolation criteria to understand how many potential minihalo candidates are excluded based on the isolation criteria and chance alignment.

The isolation of a UCHVC from existing structure depends both on its spatial and kinematic separation. For this reason, when considering the isolation of the simulated sources, we wish to account for their velocity distribution relative to

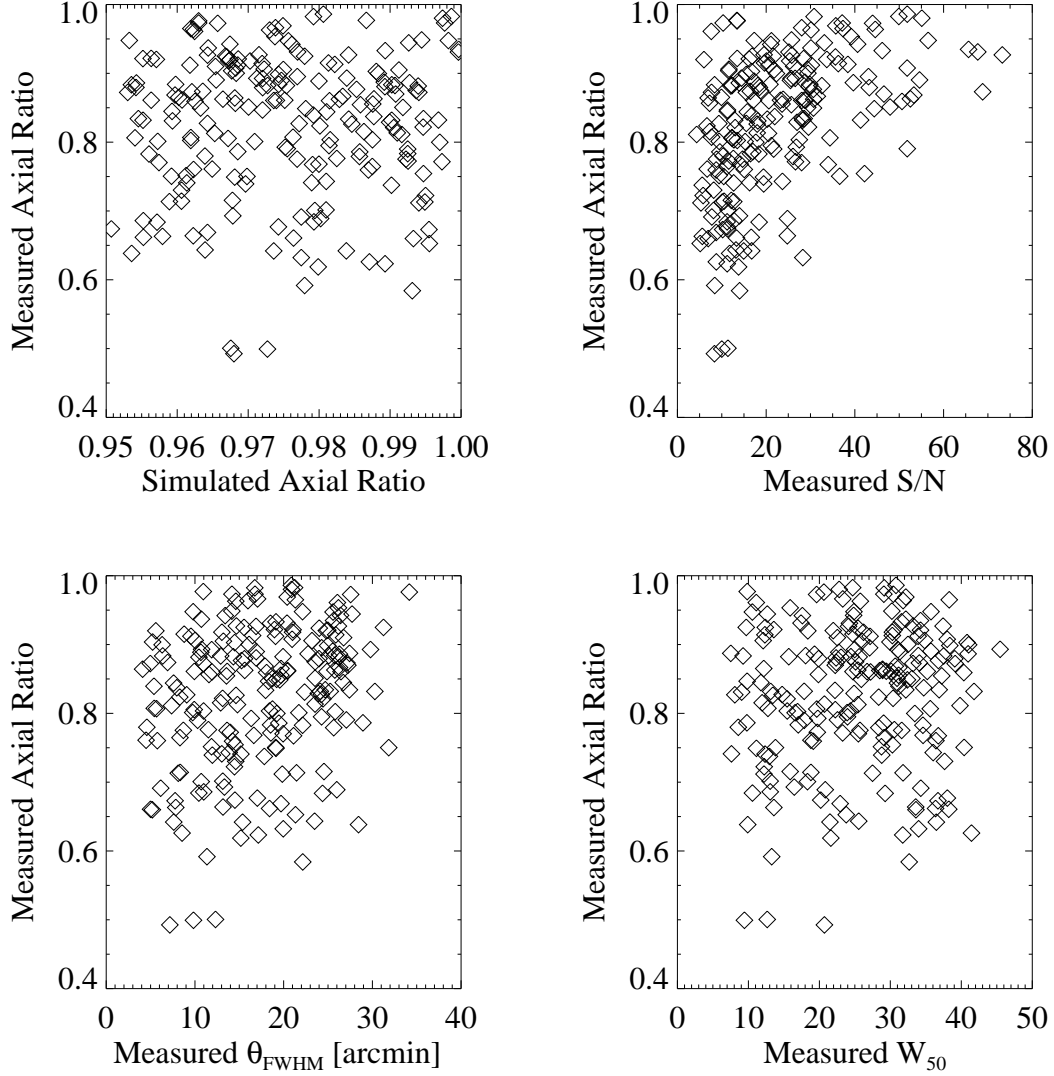


Figure 2.9: Comparison of the measured axial ratio to the simulated axial ratio (upper left). Simulated sources are nearly circular while measured sources have a range of ellipticities. The other three panels show the measured axial ratio as a function of different source parameters. High S/N sources come closest to the true axial ratio, but the measured axial ratios do not reflect the true source shape as a general rule.

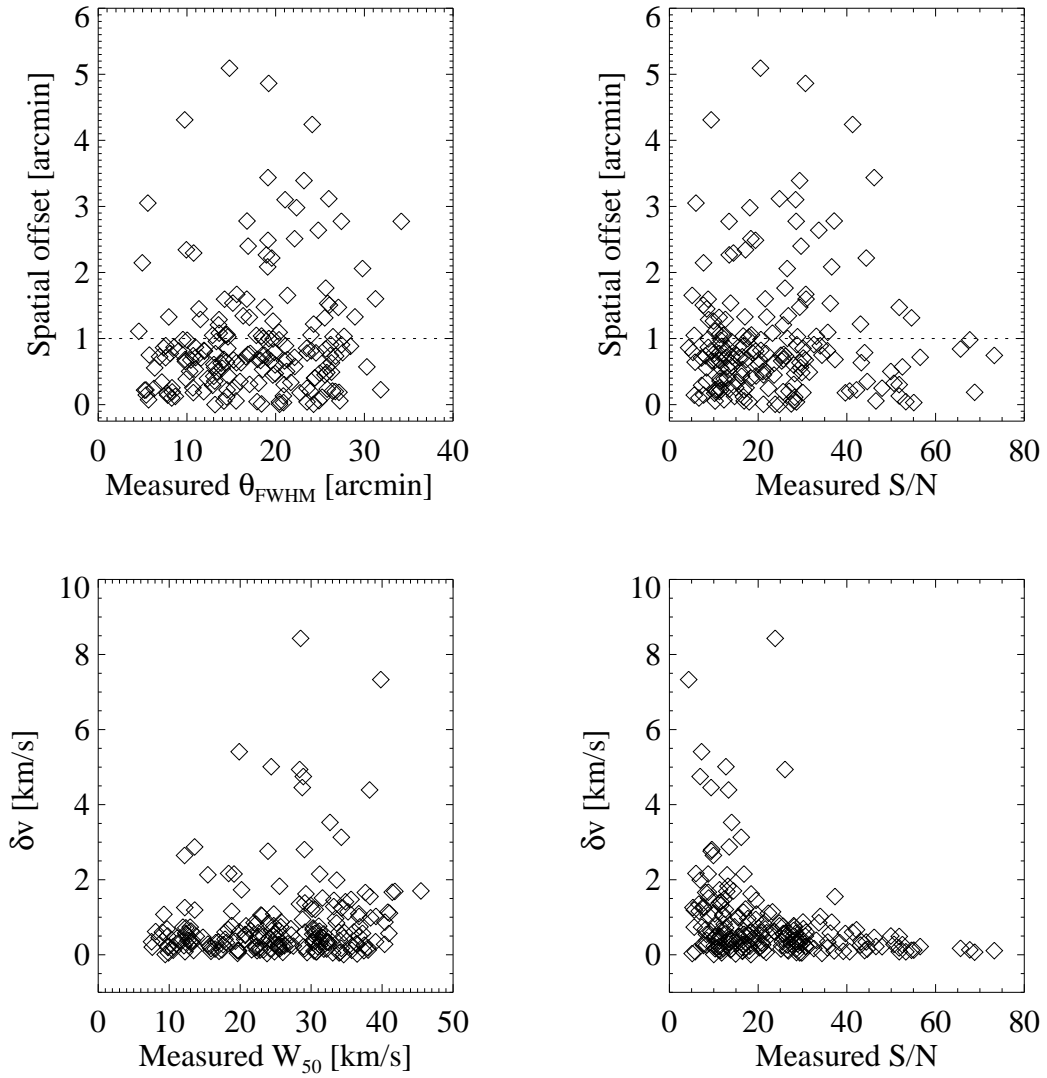


Figure 2.10: Offset of the expected source position both spatially (top row) and kinematically (bottom row) as a function of measured source parameters. Generally, the measured positions and velocities are accurate although in a few cases the offset of spatial centroid is significant.

that of HVC structure. Figure 2.11 shows the distribution of velocities for HVCs in the spring and fall sky of ALFALFA (see Section 3.1 for a discussion of what constitutes the spring and fall sky and the differences in the HVC population between the two parts of the survey). There is a clear difference in the velocity distribution between the two parts of the ALFALFA survey and from the velocity distribution of the simulated sources (see Figure 2.2 for example). For this reason, we wish to highlight the isolation of the simulated sources that lie within the velocity range of HVCs separately from the full set of simulated sources. Figure 2.11 shows the distribution of velocities for all HVCs within the spring (top panel) and fall (bottom panel) ALFALFA sky. Examining this figure, we can determine the velocity range $[-500, 200] \text{ km s}^{-1}$ to be the “HVC Velocity Range” for the fall sky, and for the spring sky we consider the velocities $[-200, 400] \text{ km s}^{-1}$ to be the “HVC Velocity Range”. Of the 300 simulated sources, 127 fall within the HVC velocity range.

First we consider the isolation of the simulated sources from the large complexes of Wakker and van Woerden [1991, WvW]. Figure 2.12 shows the isolation of the simulated sources from the large complexes; the x-axis shows the distance to the nearest WvW cloud in a complex and the y-axis shows the fraction of objects whose closest neighbor is at that distance or closer (cumulative fraction). The dotted vertical line indicates the isolation criteria defined in Chapter 3; sources to the left of this line are not considered isolated. The solid line indicate all simulated sources while the dashed line indicates the simulated sources that fall within the range of HVC velocities. The vast majority of simulated sources pass the criterion for isolation from the large HVC complexes. In fact, only 12 of the 300 simulated sources fail this criterion.

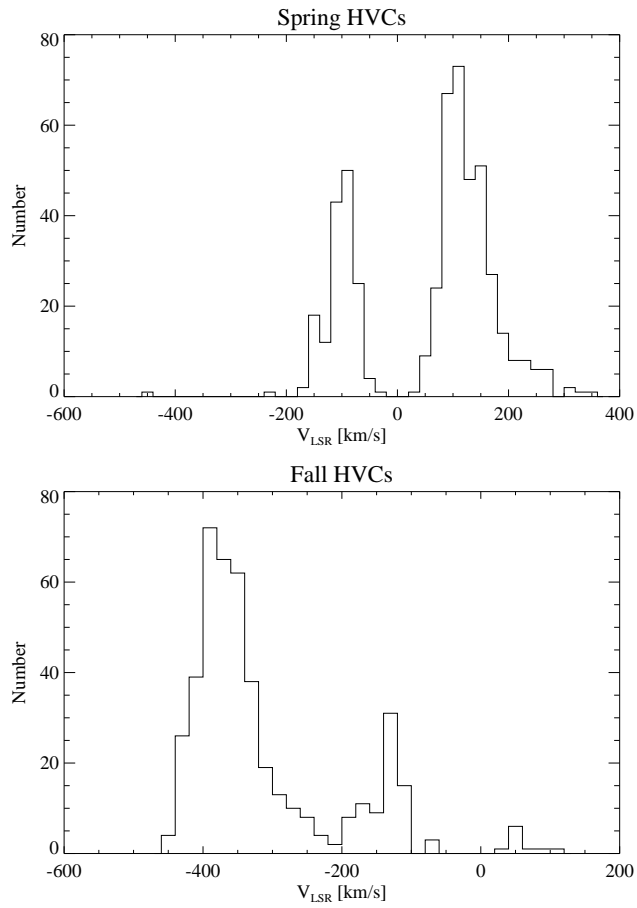


Figure 2.11: Distribution of recessional velocities for fall and spring HVCs within the ALFALFA survey. The fall HVCs occupy the velocity range $[-500, 200]$ km s^{-1} , and the spring HVCs the range $[-200, 400]$ km s^{-1} .

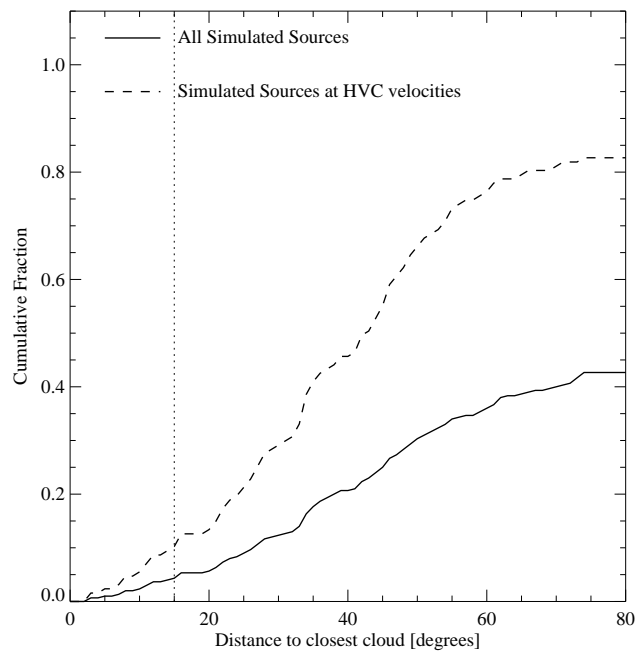


Figure 2.12: Distance of simulated sources from the nearest cloud in a WvW complex. The majority of simulated sources pass this isolation criterion.

Next we consider the isolation of simulated sources from smaller HVC structure uncovered within the ALFALFA survey. Figure 2.13 shows the isolation of the simulated sources from HVC structure within the ALFALFA survey; in the upper panel the x-axis shows the number of HVC neighbors within $D = 3^\circ$ and the y-axis is the cumulative fraction of the number of simulated sources with that many neighbors or fewer. As in the previous figure, the solid line indicates all simulated sources and the dashed line indicates the simulated sources within the range of HVC velocities. The ALFALFA isolation criterion requires no more than three neighbors within $D = 3^\circ$; the vast majority of simulated sources pass this criterion. In fact, only two simulated sources fail this criterion, and they also fail the earlier WvW isolation criterion. We can conclude that very few potential minihalo candidates are excluded from our UCHVC catalog on the basis of the isolation criteria and chance placement near HVC structure.

In the bottom panel of Figure 2.13, we examine the simulated sources in the context of the most-isolated criterion for the UCHVCs; this criterion requires sources to have no more than four neighbors out to $D = 10^\circ$ within the ALFALFA survey. As discussed in Chapter 3, this criterion selects a subsample of UCHVCs that are the most isolated and hence may represent the best minihalo candidates. As can be seen in the bottom panel of Figure 2.13, the majority of the simulated sources also pass this criterion; only 29 of the 300 simulated sources do not meet the extra criterion. This indicates that the sources that meet this extra isolation criterion are the objects independent of large-scale HVC structure and represent the best minihalo candidates.

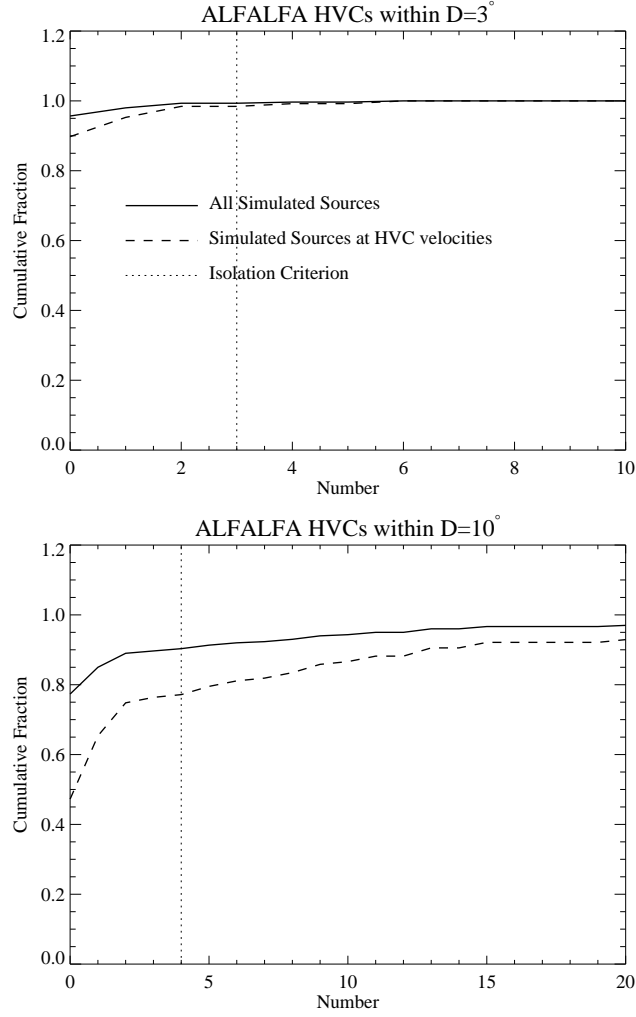


Figure 2.13: Number of ALFALFA HVC neighbors within $D = 3^\circ$ (top panel) and $D = 10^\circ$ (bottom panel). Very few sources fail the basic isolation criterion in the top panel, and the majority pass the secondary most isolated (MIS) criterion in the bottom panel, indicating that the MIS UCHVCs represent the best minihalo candidates.

2.10 Summary

I developed a 3D matched filter signal extraction algorithm designed for the narrow velocity widths and extended spatial sizes of the UCHVCs. Artificial sources in the ALFALFA data are used to test the performance of this algorithm. Generally, and as designed, this algorithm outperforms that of S07, especially for narrow velocity widths and low S/N values. While source properties reported in the catalog are measured interactively, this algorithm does return accurate source parameters within the range of templates used. Visual inspection as part of the source identification scheme increases the fraction of sources identified, especially for the narrowest velocity widths and allows sources within the Galactic hydrogen to be considered. Measured source properties (θ_{HI} , W_{50} , cz) are consistent with the simulated source values, lending credence to the catalog presented in Chapter 3. However, it should be noted that while the measured θ_{HI} values are consistent with the input source angular diameter, the ellipticity values are far from accurate; the simulated sources are nearly circular but often have measured axial ratios from 0.6 – 1. Hence, the average sizes of the UCHVCs are well constrained but not the individual axis measurements. Very few simulated sources fail the isolation criteria of Chapter 3 or the more restrictive most isolated (MIS) criterion indicating the MIS UCHVCs are indeed the best minihalo candidates.

CHAPTER 3

CATALOG OF ULTRA-COMPACT HIGH VELOCITY CLOUDS WITHIN THE $\alpha.40$ FOOTPRINT

This chapter presents a catalog of UCHVCs from the $\alpha.40$ footprint and addressed the minihalo hypothesis for these objects as a class. In Section 3.1 we discuss the $\alpha.40$ data and selection of UCHVCs. In Section 3.2 we present the UCHVC catalog and overview the observed properties of the UCHVCs. In Section 3.3 we examine the UCHVC population in the context of the known high velocity cloud (HVC) populations, and in Section 3.4 we present evidence supporting the LG origin and minihalo hypothesis for the UCHVCs. In Section 3.5, we summarize our findings.

3.1 Data

The sources presented here are found within the footprint of the $\alpha.40$ release of the ALFALFA survey [Haynes et al., 2011] but correspond to a separate analysis of the same spectral data cubes. We briefly describe the ALFALFA survey below, with an emphasis on its relevance to UCHVCs, followed by a description of how the UCHVCs are identified and measured. The ALFALFA sky is divided into two regions, termed the “spring” and “fall” as a result of our nighttime observing in the Northern Hemisphere. The “spring” ALFALFA sky covers a range of $7.5h - 16.5h$ in RA; the “fall” sky is $22h - 3h$ in RA. The $\alpha.40$ footprint covers approximately 2800 square degrees and includes the declination ranges 4° - 16° and 24° - 28° in the spring, and 14° - 16° and 24° - 32° in the fall. We note here that Leo P is located at

*This chapter is published in Adams et al. (2013)

+18° and is not in the $\alpha.40$ footprint, and hence is not included in the UCHVC sample. The footprint of the $\alpha.40$ survey can be seen in Figure 3.1; the top panel is the spring sky and the bottom panel is the fall sky. The relative sizes of the panels indicate the different RA coverage of the separate survey areas. The open diamonds in the figure show the general HVC population of the $\alpha.40$ survey and the filled symbols are the UCHVCs of this work with the color scale indicating the velocities of the clouds. The fall sky shows a prevalence of HVCs; in comparison, the spring sky is relatively clean, making this a better location to look for low mass gas-bearing dark matter halos.

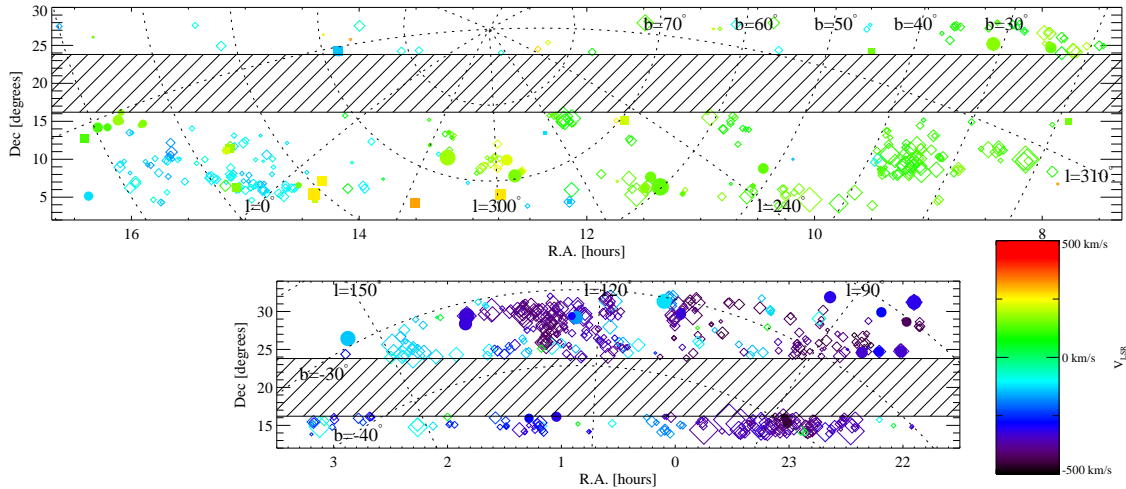


Figure 3.1: UCHVCs (filled circles) plotted in R.A.-Dec. coordinates; color corresponds to the velocity of the cloud. The solid squares are the most-isolated subsample of UCHVCs (see Section 3.1.4). The open diamonds are the $\alpha.40$ HVCs shown for reference. The size of the symbols is proportional to the angular sizes of the HVCs in all cases but not to scale. The top panel is the spring R.A. region; the bottom panel the fall R.A. region. The hashed region corresponds to declination ranges not covered by $\alpha.40$. The fall sky shows prevalent HVC structure while the spring sky is relatively clear of HVCs.

3.1.1 The ALFALFA Survey

ALFALFA is an extragalactic spectral line survey making use of the Arecibo 305m telescope. The survey maps 7000 square degrees of sky in the HI 21cm line, covering the spectral range between 1335 and 1435 MHz (roughly -2500 km s^{-1} to 17500 km s^{-1} for the HI line), with a spectral resolution of 25 kHz, or $\sim 5.5 \text{ km s}^{-1}$ (at $z = 0$). ALFALFA is designed to outperform previous blind HI surveys. With an angular resolution of $\sim 3''.5$, ALFALFA can resolve structures $1/4$ the angular size possible with the HI Parkes All Sky Survey [HIPASS; Meyer et al., 2004] and $1/9$ that possible with the Leiden Dwingeloo Survey [LDS; Hartmann and Burton, 1997]. Its flux density sensitivity is nearly one order of magnitude higher than that of HIPASS and more than two orders of magnitude better than that of the LDS. ALFALFA can detect a $\sim 5 \times 10^4 M_{\odot}$ cloud of 20 km s^{-1} linewidth at a distance of 1 Mpc. A full description of the observational mode of ALFALFA is given in Giovanelli et al. [2007], while the definition and goals of the survey are described in Giovanelli et al. [2005]. The data taking mode of ALFALFA is described briefly in §2.1.1.

3.1.2 Source Identification

As mentioned in Chapter 2, the standard source identification and measurement in ALFALFA uses the algorithm developed by Saintonge [2007] to identify sources and is then followed by measurement of the source by hand. As discussed in §2.1.3, the UCHVCs are fundamentally different from extragalactic sources due to their narrow velocity widths and spatial extent. Many of the UCHVCs are missed

by the standard identification algorithm, and a special algorithm was developed to identify the UCHVCs (presented in Chapter 2). Since the ALFALFA pipeline also involves visual inspection of the dataset, most of these sources are identified by eye and included in the $\alpha.40$ catalog as HVC detections. The specialized UCHVC identification algorithm does find sources that are missed by the standard ALFALFA pipeline; of the 59 UCHVCs identified here, 5 sources are not included in the $\alpha.40$ catalog. Three of these are in the spring sky and two in the fall sky. Figure 3.2 shows the measured properties of all the UCHVCs compared to the 5 sources not included in the $\alpha.40$ catalog. The additional sources tend to have low integrated flux densities and narrow linewidths (W_{50}). While they have a range of HI diameters, they are not the most compact clouds. Most strikingly, the UCHVCs not included in the $\alpha.40$ catalog are the sources with the lowest average column densities, suggesting that these sources are the tip of the iceberg for further clouds to be detected.

3.1.3 Criteria for UCHVC Identification

To be included as a UCHVC in the catalog, a source must have $|v_{LSR}| > 120 \text{ km s}^{-1}$, have a HI major axis less than $30'$ in size, and have a $S/N \geq 8$ to ensure reliability. The v_{LSR} limit is imposed to focus on a class of clouds that are well separated from Galactic emission and that could trace dark matter halos within the LG. Some dark matter halos would be expected to have $|v_{LSR}| < 120 \text{ km s}^{-1}$ (Leo T, for example) but disentangling their emission from Galactic hydrogen is challenging and left to future work. The $30'$ size limit corresponds to a physical size of 2 kpc at a distance of 250 kpc. The distance of 250 kpc is a reasonable minimum distance for an

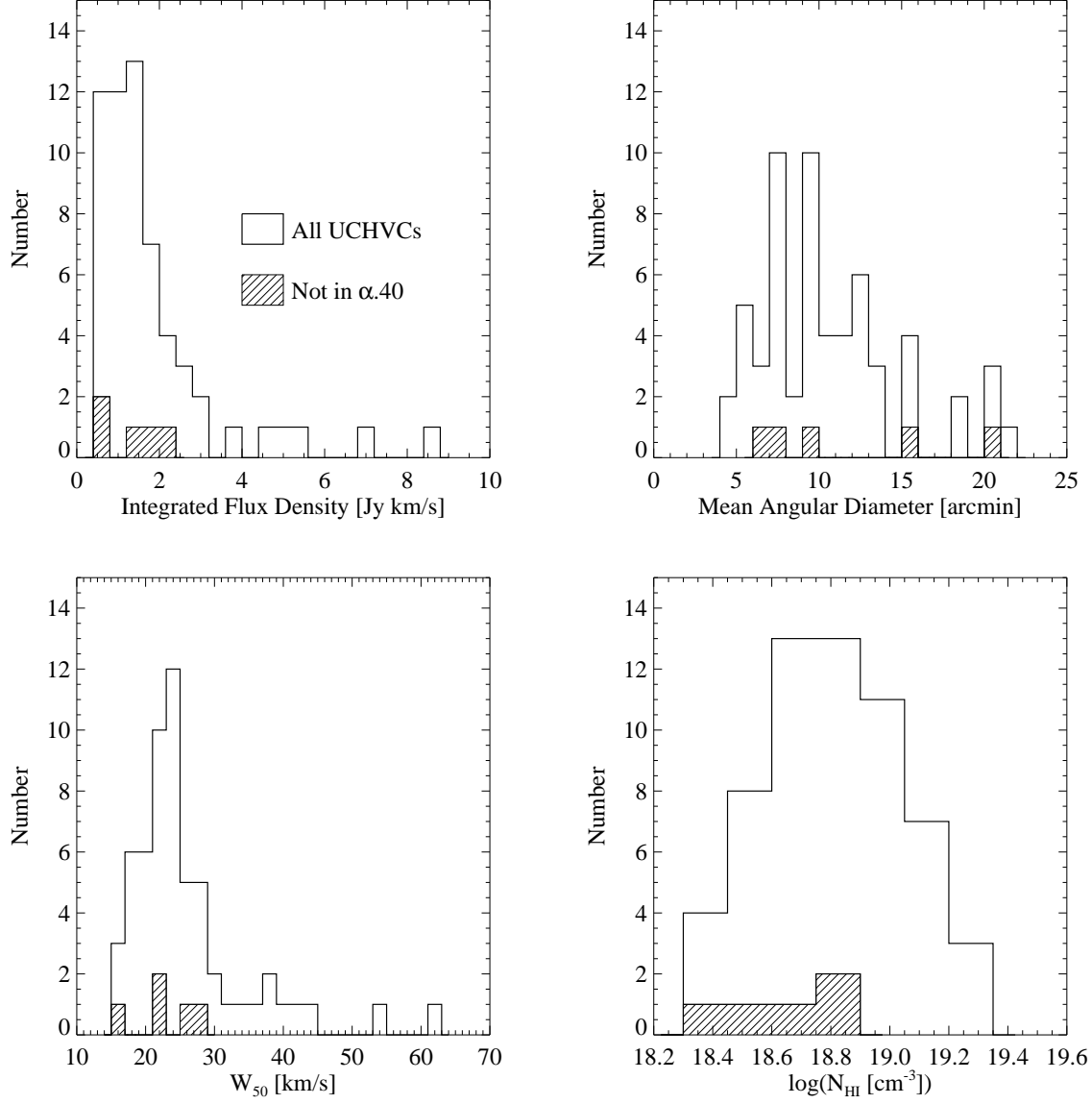


Figure 3.2: The measured properties for the full sample of UCHVCs (unfilled histograms) compared to the UCHVCs found specifically by the separate analysis presented in this work (hashed histograms). Generally, the new UCHVCs have narrow velocity widths and low fluxes. They also have low \bar{N}_{HI} values.

unperturbed object at the edge of the MW; Grcevich and Putman [2009] find that LG dwarf galaxies with neutral gas content 270 kpc away from either the MW or M31¹. We would not expect to detect low mass galaxies with large gas reservoirs nearer to the MW due to interaction with the hot Galactic corona [Fukugita and Peebles, 2006]. Note that Leo T has an HI diameter of 0.6 kpc, and Leo P has an HI diameter of 1.2 kpc. The models of Sternberg et al. [2002] for gas in dark matter minihalos predict HI diameters up to 3 kpc with diameters less than 2 kpc a more common outcome. It should be noted that most UCHVCs are smaller than this criterion, with only six clouds having average HI diameters larger than 16' (see Figure 3.2). The S/N limit of 8 ensures reliability. This limit is higher than the general reliability limit of the ALFALFA survey data due to the different nature of the UCHVCs, including the strong potential for radio frequency interference (RFI) to masquerade as narrow-line sources. Confirmation observations of low S/N sources are ongoing, and in future work we will examine the reliability and completeness of the ALFALFA UCHVC catalog.

3.1.4 Isolation

Given the abundance of HVC structures in the sky, the most important criterion for determining if a cloud is a good minihalo candidate is its isolation. Most of the known HVC structure is associated with Galactic processes, including accretion onto the Milky Way; when considering clouds that could represent gas associated with dark matter halos, we wish to find objects distinct from existing HVC struc-

¹The Magellanic Clouds do have a substantial neutral gas content and are much closer to the MW than 250 kpc. However, they are more massive than the general population of dwarfs in the LG and are actively losing their HI via interactions with the MW.

tures. In order to be considered a UCHVC, visual inspection must ensure that the cloud does not appear to be associated with a larger HI structure.

Our second isolation criterion is that the UCHVCs must be well separated from previously known HVC complexes. We compare the UCHVCs to the updated catalog of Wakker and van Woerden [1991, B. Wakker, private communication 2012; hereafter WvW]. The WvW catalog includes 617 clouds, of which 393 are classified as belonging to 20 large complexes; the other clouds are classified into populations based on their spatial coordinates and velocity. For defining isolation, we only consider the WvW clouds which are part of a larger complex. The distance of a UCHVC from another cloud in degrees can be quantified via:

$$D = \sqrt{\theta^2 + (f\delta v)^2}, \quad (3.1)$$

where θ is the angular separation in degrees, δv is the velocity difference in km s^{-1} between two clouds, and f is a conversion factor that parameterizes the significance we ascribe to the angular separation between two clouds versus their difference in velocity in determining whether they are associated with each other. Following Saul et al. [2012] and Peek et al. [2008], we adopt $f = 0.5^\circ/\text{km s}^{-1}$ as the weighting for the velocity separation for large scale HVC structure. Figure 3.3 illustrates our determination of the isolation criterion for deciding if the UCHVCs are separated from the WvW complexes. The isolation criterion was determined by comparing the separation of clouds within WvW complexes to the separation of LG galaxies from the nearest WvW cloud in a complex. The x-axis shows the distance to the nearest WvW cloud in a complex and the y-axis shows the fraction of objects whose closest neighbor is at that distance or closer (cumulative fraction). Ninety percent of WvW clouds in complexes are closer than 15° to their nearest neighbor in the complex; more than eighty percent of LG galaxies are located fur-

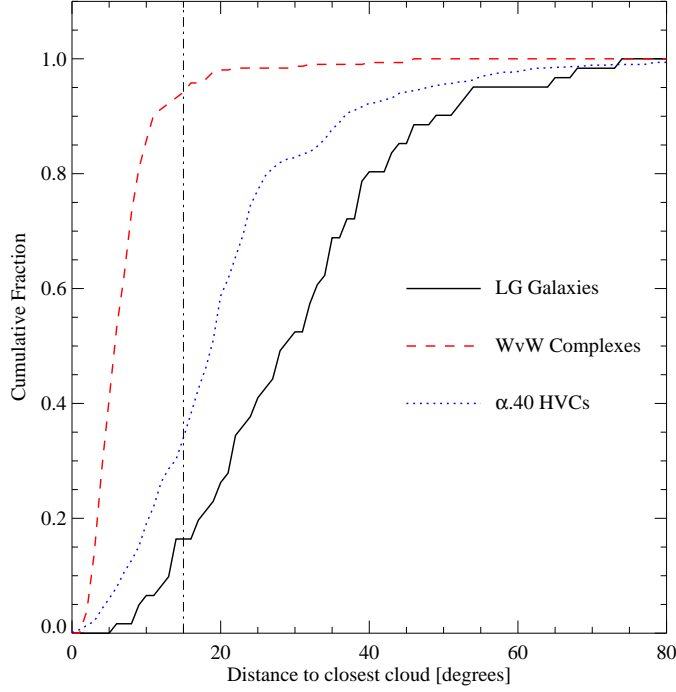


Figure 3.3: The relative isolation of LG galaxies and $\alpha.40$ HVCs from the large HVC complexes of WvW. The x-axis is the distance to the closest WvW cloud in a complex calculated using Equation 3.1. The y-axis shows the fraction of objects that have their nearest neighbor at that distance or nearer. The separation of WvW clouds within complexes from each other is shown by the dashed red line. The isolation of the LG galaxies is shown by the solid black line, and the $\alpha.40$ HVCs are shown for reference with the dotted blue line. The dot-dash line indicates our chosen isolation criterion of $D = 15^\circ$. The majority of clouds in complexes are within 15° of their nearest neighbor, although there is a smaller tail extending to 25° . The majority of LG galaxies are located further than 15° from a cloud in a complex, making this a good isolation criterion. This isolation criterion removes $\sim 30\%$ of the $\alpha.40$ HVCs from consideration as UCHVCs, but further isolation criteria are clearly necessary.

ther than 15° from the nearest WvW cloud in a complex. Hence we determine to use this value as our cutoff, shown by the dot-dash line in Figure 3.3. We note that is a more generous criteria than that of Saul et al. [2012] and Peek et al. [2008] who adopt $D = 25^\circ$ as an isolation criterion; in Section 3.3.1 we examine this intermediate distance and determine it does not substantially affect our catalog.

In addition, we institute a third isolation criterion based on HVC structure uncovered by ALFALFA. This structure is generally much smaller than previously known HVC structure; as can be seen in Figure 2.1 most $\alpha.40$ HVCs are less than one degree in size while the sizes of the HVCs in the WvW catalog are several to tens of degrees². For this reason, we use $f = 0.2^\circ/\text{km s}^{-1}$ in Equation 3.1 when calculating isolation from HVC structure within the ALFALFA survey. The top panel of Figure 3.4 shows the final isolation criterion for UCHVCs and compares the UCHVCs to LG galaxies and the general HVC detections within the $\alpha.40$ survey. We require that the UCHVCs have no more than three neighbors within $D = 3^\circ$. This is a generous criterion as the LG galaxies have at most one neighbor within this distance. We wish to include all potential minihalo candidates and inspection indicates that allowing three neighbors includes all the sources that would be classified by eye as isolated. In the bottom panel of Figure 3.4 we explore the differences between the spring and fall populations of the UCHVCs. The fall sky appears to show more isolation on this scale with the UCHVCs having either one or no neighbors; in fact, this is a result of the prominent HVC structure in the fall sky. Clouds in the fall sky are either part of a larger structure or have no (or one) neighbors within $D = 3^\circ$. Comparing to the general $\alpha.40$ HVC population

²As an extragalactic survey, ALFALFA was not designed to detect sources with sizes $\gtrsim 1^\circ$; the commensal GALFA-HI survey which processes the signal independently does that [e.g. Peek et al., 2011].

shows the prevalence of HVC structure in the fall sky with the fall HVCs generally having more neighbors than the spring HVCs.

We note that with this criterion, only clouds with central velocities within 15 km s^{-1} of the UCHVC can be considered as neighbors. Given that the median velocity width of the UCHVCS is 23 km s^{-1} , there is a possibility that this isolation criterion could leave our sources kinematically confused. Our first isolation criterion accounts for this through the examination of the UCHVCs for association with other clouds. In order to verify this, we examine the effect of changing the velocity weighting factor to $f = 0.05^\circ / \text{km s}^{-1}$. This expands the velocity selection to 60 km s^{-1} , almost three times the median FWHM of the clouds. We examine the number of clouds within 3° of the UCHVCs using this different value of f and find that the UCHVCs still have very few neighbors with this modified distance estimate. In fact, seventy-five percent of the UCHVCs still meet the criterion of three or fewer neighbors even when the expanded velocity space is considered. We examined the nine UCHVCs with more than five neighbors and note that three of them may possibly be kinematically associated with larger structure.

HVC structure often exists on scales much larger than 3° ; while the UCHVCs are examined for obvious connection to larger structure and excluded in that case, we still wish to define a more isolated subsample. As the best subsample to represent HI sources associated with minihalo candidates, we define a “most-isolated” subsample (MIS) of UCHVCs with no more than 4 neighbors within $D = 10^\circ$. The top panel of Figure 3.5 shows the number of neighboring clouds within $D = 10^\circ$ for the UCHVCs, the MIS UCHVCs, LG galaxies and $\alpha.40$ HVCs. On this large scale, the MIS UCHVCs are generally more isolated than even the LG galaxies. We do note that the $\alpha.40$ footprint means that we are not generally

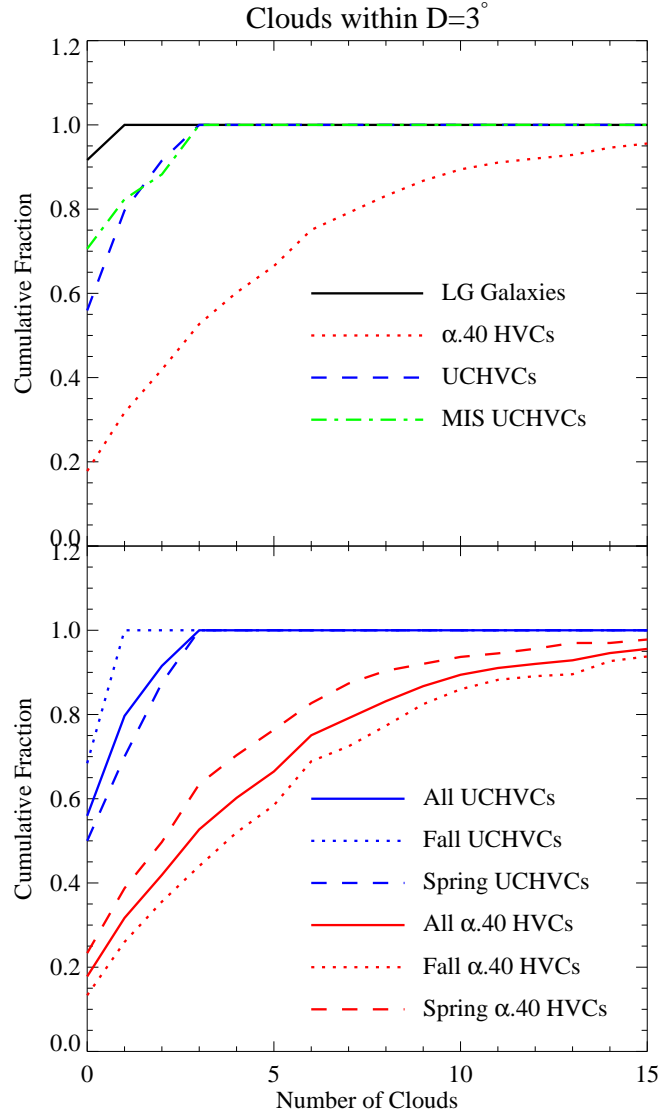


Figure 3.4: The x-axis is the number of $\alpha.40$ HVCs within $D = 3^\circ$, where the distance is calculated from Equation 3.1 with $f = 0.2^\circ/\text{km s}^{-1}$, and the y-axis is the fraction of UCHVCs with that number of neighbors or fewer. The top panel shows the relative isolation of LG galaxies (solid black line), UCHVCs (dashed blue line), MIS UCHVCs (dot-dash green line), and general $\alpha.40$ HVCs (dotted red line). The LG galaxies have no more than one $\alpha.40$ HVC within $D = 3^\circ$; the criteria for the UCHVCs is slightly relaxed to not more than 3 neighbors. The $\alpha.40$ HVCs are shown for reference; a majority of the $\alpha.40$ HVCs fail this isolation criteria. In the bottom panel, we compare the spring (dashed line) and fall populations (dotted line) of the UCHVCs (blue), with the $\alpha.40$ HVCs shown for references (red).

probing to a full 10° in all directions around a given cloud; increasing coverage of the ALFALFA survey may change the classification of a cloud in the future. In fact, two sources in the fall $\delta = +15^\circ$ strip meet the MIS criteria but we exclude them from this subsample as determining isolation out to 10° for sources in an isolated 2° wide strip is problematic. We will revisit these two specific sources and the classification of the MIS UCHVCs in general with increased ALFALFA coverage in future work. In the bottom panel of Figure 3.5, we again examine the difference between the fall and spring population. Here, the prominent HVC structure in the fall sky is apparent with many of the fall UCHVCs having a large number of neighbors out to a distance of 10° . There is also a strong difference evident between the UCHVC and $\alpha.40$ HVC population with over half of the $\alpha.40$ HVCs having more than 20 neighbors at $D = 10^\circ$; this indicates the utility of our first isolation criterion of inspecting sources for connection to large scale structure.

3.2 Catalog

3.2.1 Presentation of Catalog

In Table 3.1 we present the UCHVCs; there are 59 sources total: 40 in the spring $\alpha.40$ sky and 19 in the fall sky. Of the 59 UCHVCs, 17 are identified as being in the most-isolated subsample, all of which are in the spring sky. The spring sky samples the outer regions of the LG where the expected density for dark matter halos may be lower but the environment is safer for gas-bearing minihalos than near the MW or M31. The fall sky samples the LG near M31 and includes the presence of a large amount of HVC structure, including the Magellanic Stream (see

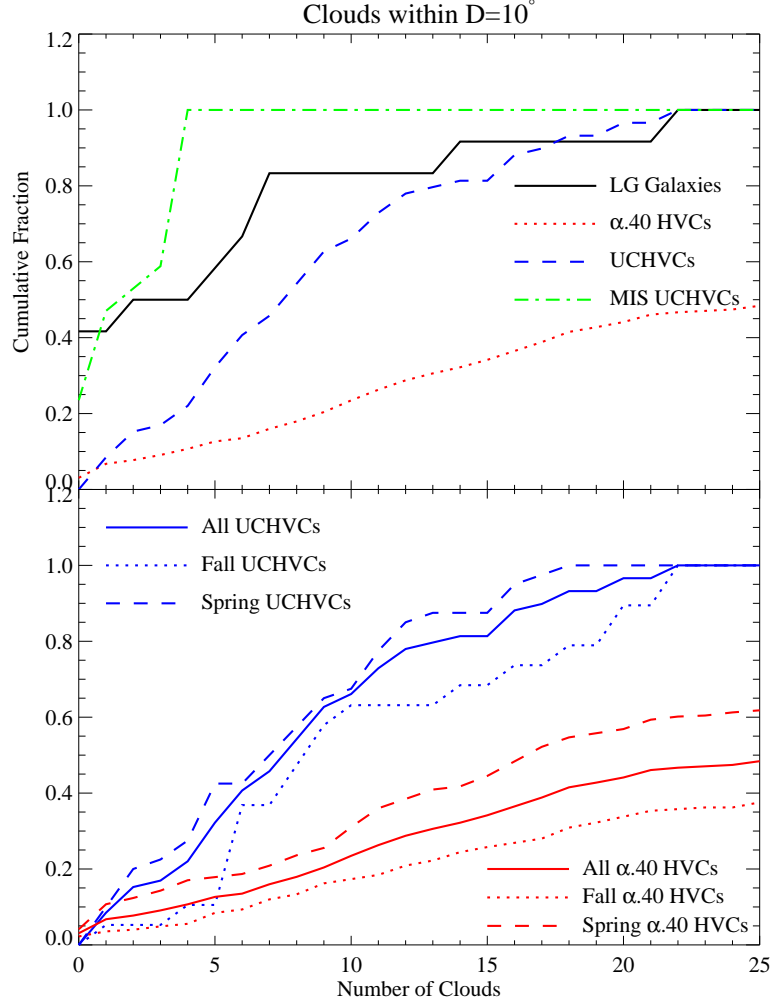


Figure 3.5: The x-axis is the number of $\alpha.40$ HVCs within $D = 10^\circ$, where the distance is calculated from Equation 3.1 with $f = 0.2^\circ/\text{km s}^{-1}$, and the y-axis is the fraction of UCHVCs with that number of neighbors or fewer. The top panel shows the relative isolation at this larger distance scale of LG galaxies (solid black line), UCHVCs (dashed blue line), MIS UCHVCs (dot-dash green line), and general $\alpha.40$ HVCs (dotted red line). At this distance scale, the UCHVCs and LG galaxies have similar behavior. We define a most-isolated subsample (MIS) of UCHVCs which are still isolated with no more than 3 neighbors on this larger scale. The MIS UCHVCs are even more isolated than the LG galaxies on this larger scale. In the bottom panel, we compare the spring (dashed line) and fall populations (dotted line) of the UCHVCs (blue), with the $\alpha.40$ HVCs shown for references (red).

Section 3.3.1 for a further discussion). We indicate those UCHVCs that are part of the original sample of UCHVCs discussed by G10 with a G10 in the notes column and those UCHVCs that lie outside the area considered by G10 with an ‘O’. Figure 3.6 shows maps of all the UCHVCs with contours in units of column density of HI (N_{HI} in atoms cm^{-2}), representing the sum total of HI content along the line of sight; these plots represent the data from which all the parameters listed in Table 3.1 are derived. The minimum contour level is given in the figure and subsequent contour levels increase by factors of $\sqrt{2}$. We plot the contours in values of N_{HI} to demonstrate that the peak column density value is higher than the average value calculated later (see Section 3.2.4). However, we emphasize, that since these clouds are barely resolved by the Arecibo beam, the column density contour values are only approximate and the average values are more robust; to accurately map the distribution of HI will require synthesis observations that provide a smaller beam. Column density values can be derived from the brightness temperature via:

$$N_{HI} = 1.823 \times 10^{18} \int T_B dv \text{ [cm}^{-2}\text{]}. \quad (3.2)$$

In simple cases, the brightness temperature is related to the flux density at 21cm via:

$$T_B = \frac{606}{\theta^2} S \quad (3.3)$$

where θ is the (circular) beam in arcseconds and S the flux in mJy/beam.

The columns of the tables are as follows:

- Col. 1: Source name, in the traditional form for HVCs, obtained from the galactic coordinates at the nominal cloud center and the v_{LSR} of the cloud, e.g. HVC111.65-30.53-124 has $l = 111.65^\circ$, $b = -30.53^\circ$, and $v_{LSR} = -124 \text{ km s}^{-1}$.

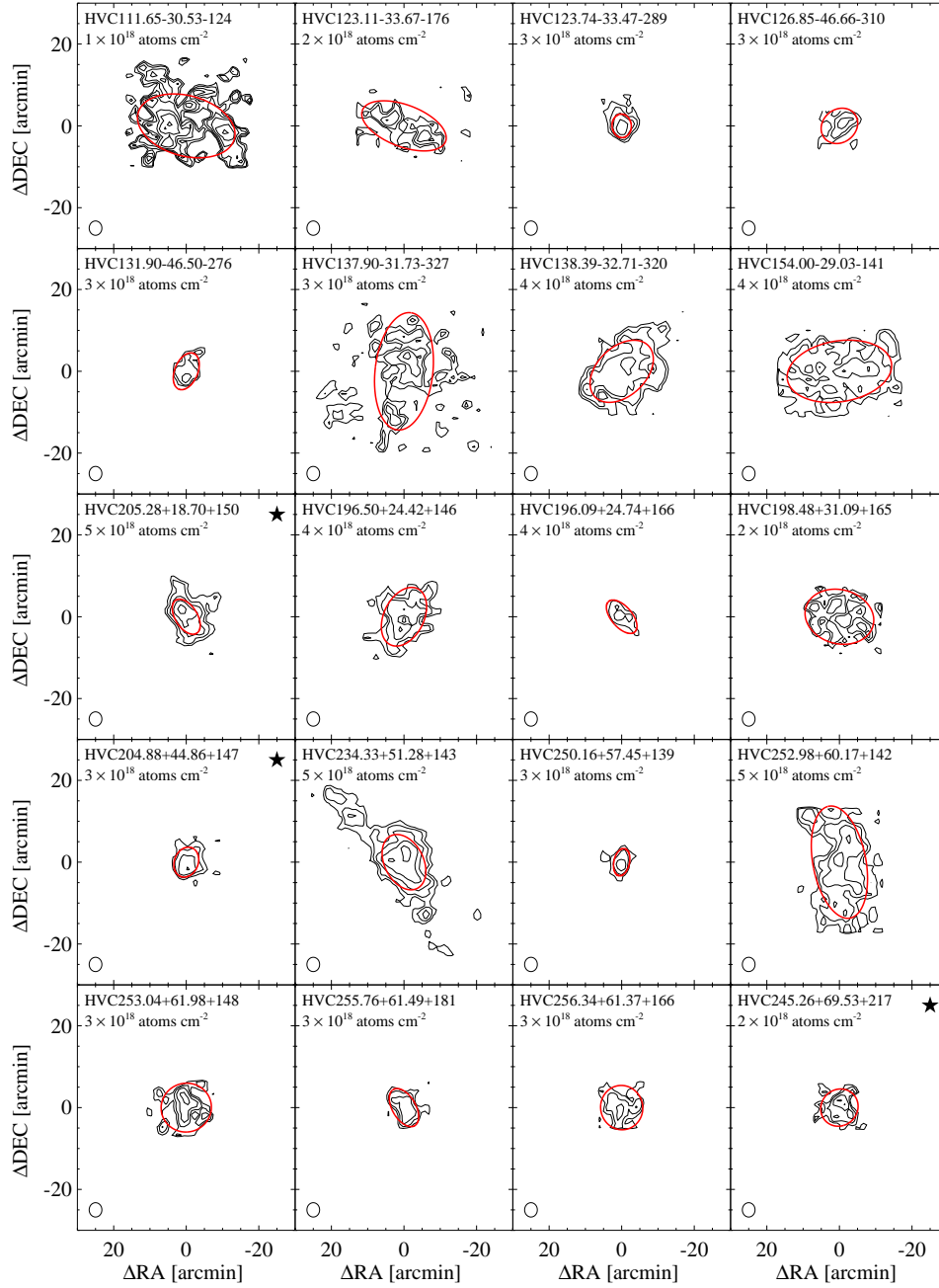


Figure 3.6: Maps of the HI column density of the UCHVCs derived from ALFALFA spectral grids. Starred figures indicate membership in the most-isolated subsample. Ellipses (red) represent the measured half-power level. The $3'.5$ circular beam is shown in the lower left corner of all plots. The lowest contour level is listed in the upper left corner of each plot; subsequent contours increase by factors of $\sqrt{2}$.

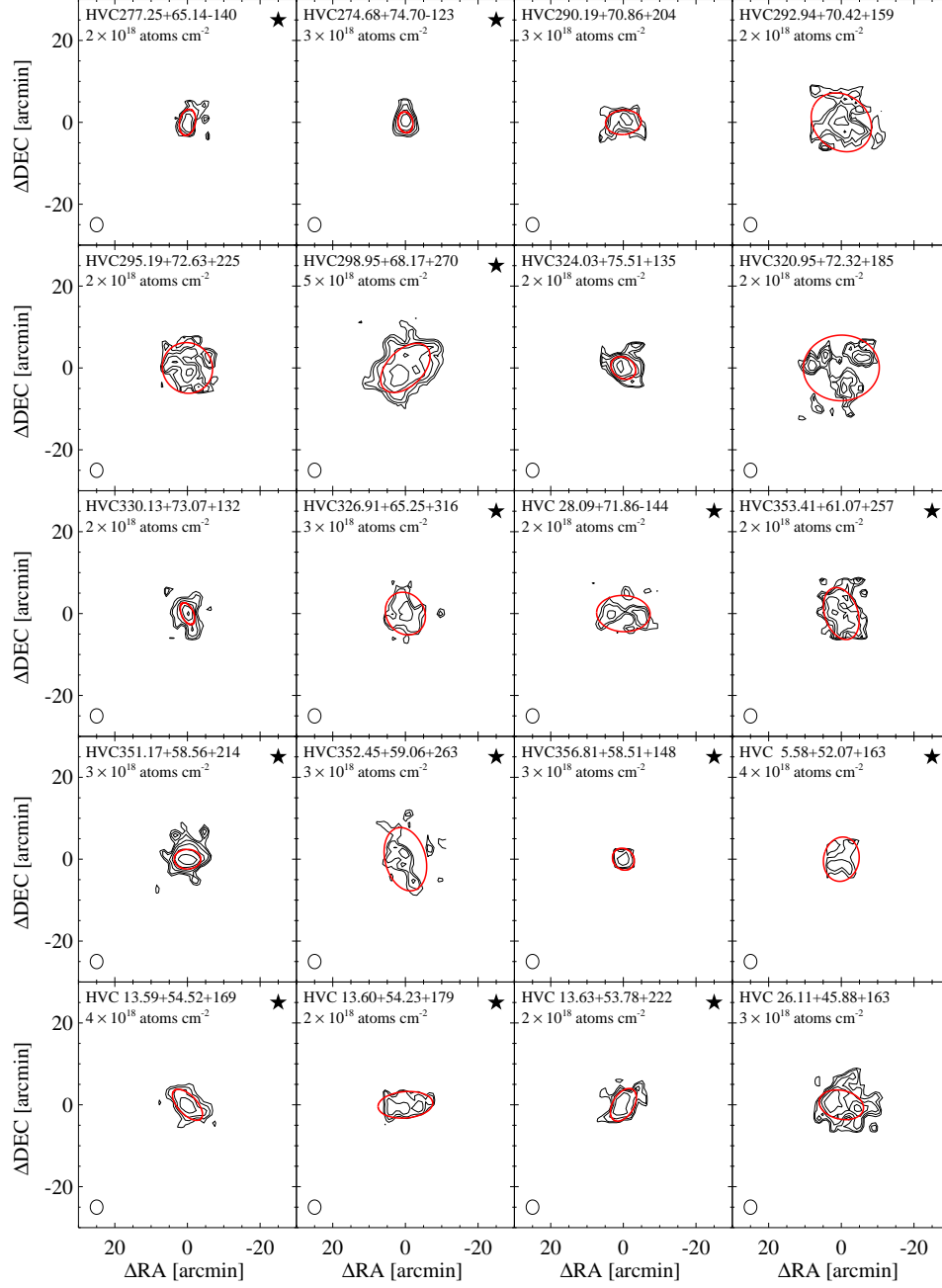


Figure 3.6: Continued

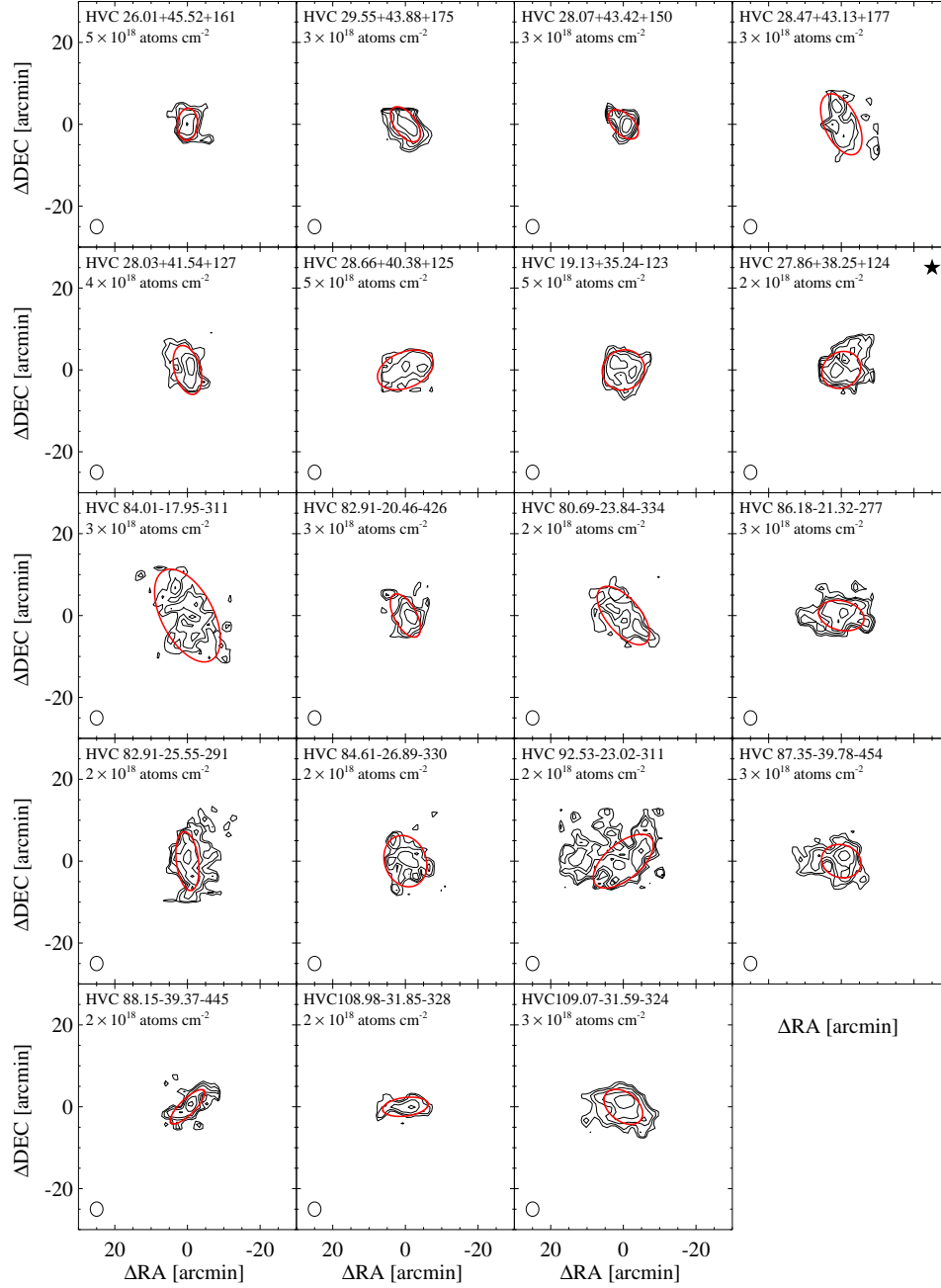


Figure 3.6: Continued

- Col. 2: Identification number in the Arecibo General Catalog (AGC), an internal database maintained by MH and RG, included to ease cross-reference with our archival system and the α .40 catalog.
- Col. 3: Equatorial coordinates of the centroid, epoch J2000. Typical errors are less than $1'$.
- Col. 4: Sequentially, we list heliocentric velocity, velocity in the local standard of rest frame (LSR; assumed solar motion of 20 km s^{-1} towards $l = 57^\circ$, $b = 25^\circ$), velocity in the Galactic standard of rest frame (GSR; $V_{gsr} = V_{lsr} + 225 \sin l \cos b$, with both velocities in km s^{-1}), and the velocity with respect to the LG reference frame from Karachentsev and Makarov [1996].
- Col. 5: HI line full width at half maximum (W_{50}), with estimated measurement error in brackets. The notes column indicates the method of measurement: a gaussian fit or linear single peaks fit to the sides of the profile.
- Col. 6: Estimate of the cloud major and minor diameters, in arcminutes. Sizes are measured at approximately the level encircling half the total flux density. In many cases, the outer contours are more elongated than indicated by the ratio $a \times b$. The half-power ellipses are also shown in the HI column density contour plots in Figure 3.6.
- Col. 7: Flux density integral (S_{21}), in Jy km s^{-1} .
- Col. 8: Signal-to-noise ratio (S/N) of the line, defined as

$$S/N = \left(\frac{1000 S_{21}}{W_{50}} \right) \frac{w_{smo}^{1/2}}{\sigma_{rms}}, \quad (3.4)$$

where S_{21} is the integrated flux density in Jy km s^{-1} , as listed in Column 7; the ratio $1000 S_{21}/W_{50}$ is the mean flux density across the feature in mJy ;

w_{smo} is $W_{50}/(2 \times 10)$, a smoothing width, and σ_{rms} is the rms noise figure across the spectrum measured in mJy. More details on the S/N calculation are available in Haynes et al. [2011].

- Col. 9: The number of α .40 HVC neighbors within $D = 3^\circ$ (for $f = 0.2^\circ/\text{km s}^{-1}$)
- Col. 10: The number of α .40 HVC neighbors within $D = 10^\circ$ (for $f = 0.2^\circ/\text{km s}^{-1}$)
- Col. 11: Notes column. For each source there is either a ‘g’ or ‘p’ indicating the method used (gaussian or single peaks fit) to measure W_{50} . Sources considered by G10 are indicated with a ‘G10’ in the notes column. Sources that are outside the footprint considered in G10 are marked with a ‘O’. The UCHVCs that are also in the GALFA compact cloud catalog of Saul et al. [2012] are indicated with a ‘S12’.

Table 3.1. ALFALFA UCHVCs in the $\alpha 40$ Survey

Source	AGC	R.A.+Dec. J2000	cz_{\odot}	$V_{l_{sr}}$ km s^{-1}	$V_{g_{sr}}$ km s^{-1}	V_{LG}	$W_{50}(\epsilon_w)$ km s^{-1}	$a \times b$ '	S_{21} Jy km s^{-1}	S/N	N_3	N_{10}	Notes
HVC111.65-30.53-124 ^a	103417	000554.3+312014	-128	-124	55	139	21 (8)	27×15	2.31	12	0	9	g, O
HVC123.11-33.67-176	102992	005206.2+291204	-177	-176	-19	61	21 (3)	24×10	1.28	9	0	5	g, O
HVC123.74-33.47-289 ^c	102994	005431.6+292402	-290	-289	-133	-52	21 (1)	6×5	0.67	15	0	13	g, O
HVC126.85-46.66-310	749141	010237.8+160752	-308	-310	-186	-112	23 (6)	10×8	0.81	9	1	21	g, O
HVC131.90-46.50-276 ^{a,c}	114574	011703.4+155548	-273	-276	-160	-88	27 (4)	10×6	0.71	9	1	22	g, O
HVC137.90-31.73-327	114116	014952.1+292600	-325	-327	-199	-124	34 (8)	29×16	3.93	13	1	9	g, O
HVC138.39-32.71-320	114117	015031.4+282259	-317	-320	-194	-119	22 (2)	19×13	4.41	30	1	7	g, O, S12
HVC154.00-29.03-141	122836	025229.7+262630	-135	-141	-55	8	27 (3)	29×15	6.90	31	0	15	g, O
HVC205.28+18.70+150★	174540	074559.9+145837	162	150	59	42	23 (4)	10×6	2.06	28	0	2	g, S12, O
HVC196.50+24.42+146	174763	075527.1+244143	156	146	88	79	20 (2)	16×11	2.80	20	3	11	g, S12, O
HVC196.09+24.74+166	174764	075614.8+250900	175	166	110	101	24 (6)	10×5	0.66	9	3	10	p, O
HVC198.48+31.09+165	189054	082546.7+251128	173	165	104	90	26 (1)	19×13	1.77	13	0	8	g, O
HVC204.88+44.86+147★	198511	093013.2+241217	152	147	80	53	15 (1)	8×6	0.73	14	0	0	g, S12, O
HVC234.33+51.28+143	208315	102701.1+084708	148	143	29	-22	20 (2)	15×10	4.96	35	0	16	g, S12
HVC250.16+57.45+139	219214	110929.8+052601	142	139	25	-32	20 (5)	7×4	0.56	10	0	9	g, G10, S12
HVC252.98+60.17+142	219274	112119.6+062132	143	142	35	-22	27 (5)	28×15	8.55	37	1	10	g, S12
HVC253.04+61.98+148	219276	112624.8+073915	149	148	47	-8	36 (1)	14×12	2.06	14	1	11	g
HVC255.76+61.49+181	219278	112855.6+062529	182	181	77	19	18 (2)	11×6	0.90	13	0	7	g, S12
HVC256.34+61.37+166 ^c	219279	112928.6+060923	167	166	61	3	24 (1)	12×11	1.49	14	2	11	g
HVC245.26+69.53+217★	215417	114008.1+150644	216	217	146	97	17 (4)	10×9	0.70	9	0	1	g, G10
HVC277.25+65.14-140★	227977	120920.0+042330	-142	-140	-234	-294	23 (1)	7×4	0.46	8	0	1	g, G10
HVC274.68+74.70-123★	226067	122154.7+132810	-128	-123	-182	-232	54 (13)	5×4	0.92	11	0	0	p, G10
HVC290.19+70.86+204	226165	123440.2+082408	200	204	135	80	21 (1)	10×6	0.90	11	1	15	g
HVC292.94+70.42+159 ^a	229344	123758.5+074849	154	159	89	34	15 (4)	17×14	1.67	13	0	18	g
HVC295.19+72.63+225	226170	124204.6+095405	220	225	164	112	28 (7)	14×12	1.17	10	3	16	p, G10
HVC298.95+68.17+270★	227987	124529.8+052023	265	270	196	139	26 (1)	16×9	5.58	44	0	4	g, G10
HVC324.03+75.51+135	233763	131242.3+133046	127	135	102	56	29 (1)	7×5	0.94	18	1	12	g

Table 3.1 (cont'd)

Source	AGC	R.A.+Dec. J2000	cz_{\odot}	V_{lsr} km s^{-1}	V_{gsr} km s^{-1}	V_{LG}	$W_{50}(\epsilon_w)$ km s^{-1}	$a \times b$ '	S_{21} Jy km s^{-1}	S/N	N_3	N_{10}	Notes
HVC320.95+72.32+185	233830	131321.5+101257	177	185	141	92	23 (9)	21×16	1.70	9	0	15	g, G10
HVC330.13+73.07+132	233831	132241.6+115231	124	132	100	53	16 (1)	6×3	0.63	11	0	11	g, G10
HVC326.91+65.25+316★	238713	133043.8+041338	308	316	264	210	26 (4)	12×10	1.25	11	0	0	p, G10
HVC 28.09+71.86-144★	249393	141058.1+241204	-157	-144	-111	-136	43 (6)	15×9	1.12	8	0	0	g, O
HVC353.41+61.07+257★	249323	141948.6+071115	246	257	244	201	20 (4)	13×9	1.34	13	3	4	g, G10
HVC351.17+58.56+214★, ^b	249282	142321.2+043437	203	214	196	151	40 (8)	7×5	1.45	17	0	4	p, G10, S12
HVC352.45+59.06+263★	249283	142357.7+052340	252	263	248	203	32 (9)	16×11	1.11	8	3	4	g, G10
HVC356.81+58.51+148★	249326	143158.8+063520	136	148	141	100	38 (11)	6×5	0.70	10	0	1	p
HVC 5.58+52.07+163★	258459	150441.3+061259	149	163	176	141	24 (8)	11×10	1.33	13	0	4	g
HVC 13.59+54.52+169★	258237	150723.0+113256	155	169	200	170	23 (3)	10×5	1.34	17	1	3	g
HVC 13.60+54.23+179★	258241	150824.4+112422	164	179	210	180	17 (1)	15×7	0.99	11	1	4	g
HVC 13.63+53.78+222★	258242	151000.6+111127	207	222	253	224	21 (2)	9×6	0.71	9	0	1	g, G10
HVC 26.11+45.88+163	257994	155354.0+144148	146	163	232	217	23 (3)	12×7	2.04	22	2	8	g
HVC 26.01+45.52+161	257956	155507.5+142929	144	161	230	215	25 (6)	8×6	1.54	14	2	8	g
HVC 29.55+43.88+175	268067	160529.4+160912	158	175	255	244	37 (11)	10×6	1.91	20	2	6	g, G10
HVC 28.07+43.42+150	268069	160532.6+145920	132	150	227	214	29 (4)	10×5	1.15	11	0	10	g, G10
HVC 28.47+43.13+177	268070	160707.0+150831	160	177	255	243	20 (3)	17×9	1.48	11	2	6	g, G10
HVC 28.03+41.54+127	268071	161236.8+141226	109	127	206	194	62 (15)	12×7	2.67	18	1	8	g
HVC 28.66+40.38+125	268072	161745.3+141036	108	125	208	197	42 (5)	16×9	3.17	21	3	7	g
HVC 19.13+35.24-123	268213	162235.7+050848	-139	-123	-63	-81	17 (1)	12×10	2.83	22	0	7	g, G10, S12
HVC 27.86+38.25+124★	268074	162443.4+124412	107	124	207	197	23 (4)	11×9	1.28	13	2	4	g
HVC 84.01-17.95-311	310851	215406.2+311249	-324	-311	-98	-21	21 (4)	26×14	2.60	17	0	5	g
HVC 82.91-20.46-426	310865	215802.9+283735	-439	-426	-217	140	22 (1)	12×6	0.99	10	0	17	g, S12
HVC 80.69-23.84-334	321318	220100.7+244404	-345	-334	-131	-55	23 (1)	18×9	1.47	13	0	5	g
HVC 86.18-21.32-277	321455	221121.8+295402	-288	-277	-68	10	17 (1)	13×7	1.76	15	0	5	g, O
HVC 82.91-25.55-291	321320	221238.6+244311	-302	-291	-90	-13	24 (2)	15×6	1.31	13	0	7	g, O
HVC 84.61-26.89-330	321351	222134.4+243638	-341	-330	-130	-53	21 (4)	13×11	1.03	9	0	8	g, O

3.2.2 Comparison to G10

For completeness, we include in Table 3.2 the UCHVCs that were considered by G10 but do not meet the stricter selection criteria used here. The clouds from G10 can fail any of the criteria: S/N, isolation or v_{LSR} limits. The notes column indicates the reason a G10 cloud is not included here. The sources with $S/N < 8$ will be considered in future work when we extend the UCHVC catalog to lower S/N values after assessing reliability and completeness. In addition, we will extend the catalog to velocities including the Galactic hydrogen. It should be noted that the three sources that do not meet the isolation criteria only barely fail. Two sources have one and two more neighbors than allowed, respectively, and the third source is excluded based on examination of large scale structure. These sources could still be good minihalo candidates.

3.2.3 Properties of the UCHVCs

Figure 3.7 shows the distribution of measured properties for the $\alpha.40$ UCHVCs and the most-isolated subsample: integrated flux density (S_{21}), average angular diameter ($\bar{a} = \sqrt{ab}$), velocity FWHM (W_{50}), and v_{LSR} . The UCHVCs have integrated flux densities of ~ 0.66 - 8.55 Jy km s $^{-1}$, with the vast majority having integrated flux densities below 3.5 Jy km s $^{-1}$ and a median flux density of 1.34 Jy km s $^{-1}$. The singly hatched histograms are the UCHVCs in the most-isolated subsample. Note that the range of values for the MIS UCHVCs is similar to the larger UCHVC population, and the median values are essentially identical. The UCHVCs range in average diameter from essentially unresolved ($\sim 4'$) to just over

Table 3.1 (cont'd)

Source	AGC	R.A.+Dec. J2000	cz_{\odot}	V_{lsr}	V_{gsr}	V_{LG}	$W_{50}(\epsilon_w)$ km s^{-1}	$a \times b$ '	S_{21} Jy km s^{-1}	S/N	N_3	N_{10}	Notes
HVC 92.53-23.02-311	321457	223823.4+315257	-321	-311	-104	-23	28 (2)	19×9	1.68	12	0	5	g, O
HVC 87.35-39.78-454 ^a	334256	230056.4+152014	-461	-454	-282	-206	26 (4)	11×8	1.57	16	0	1	g, O
HVC 88.15-39.37-445 ^a	334257	230211.3+160048	-452	-445	-271	-195	22 (11)	12×4	0.68	10	0	4	g, O
HVC108.98-31.85-328	333613	235658.8+293235	-333	-328	-147	-64	19 (2)	13×5	0.55	8	1	19	g, O
HVC109.07-31.59-324	333494	235702.1+294846	-329	-324	-143	-60	17 (5)	12×7	1.80	23	1	19	g, O

★Part of the extremely isolated MIS subsample

^aNot included in the α .40 catalog

^bAlso included in the compact cloud catalog of Saul et al. [2012]

^cPossible kinematic association with larger structure

Table 3.2. UCHVCs from G10 that Fail UCHVC Criteria

Source	AGC	R.A.+ Dec. J2000	cz_{\odot} V_{lsr} V_{gsr} V_{LG} km s^{-1}	$W_{50}(\epsilon_w)$ km s^{-1}	$a \times b$ '	S_{21} Jy km s^{-1}	S/N	N_3	N_{10}	Reason
HVC244.51+53.41+160	208424	104850.1+050419	164 160 39 -18	19 (3)	16×12	1.03	7	0	9	S/N
HVC249.03+57.58+178	219213	110813.6+055725	179 176 64 5	19 (2)	12×9	0.67	7	0	8	S/N
HVC247.19+70.29+247	215418	114418.2+150509	246 247 177 129	30 (10)	10×8	0.54	7	0	1	S/N
HVC290.37+66.23-115	227983	123116.7+035044	-118 -114 -199 -259	20 (5)	6×4	0.44	9	0	3	Velocity
HVC298.30+72.91+185	226171	124557.2+100518	180 185 127 75	25 (3)	5×4	0.57	9	5	21	Isolation
HVC299.62+67.65+326	227988	124619.1+044923	323 327 253 195	39 (13)	14×7	0.76	6	0	0	S/N
HVC314.57+74.80+218	238626	130351.1+121223	211 218 176 127	36 (13)	5×3	0.35	5	0	17	S/N
HVC 8.88+62.16+281	249538	143531.7+133126	269 282 298 264	18 (6)	4×3	0.22	4	0	4	S/N
HVC 7.64+57.83-128	249248	144844.6+103510	-142 -128 -112 -147	22 (1)	25×5	1.83	16	0	42	Isolation
HVC 15.11+45.54-148	258474	154035.2+074334	-163 -147 -106 -132	27 (1)	7×5	0.68	9	4	19	Isolation

20' in size, with the vast majority less than 16' in size and a median size of 10'. We note that there does appear to be a break in population based on size with UCHVCs clustered with HI diameters $< 16'$ in size and a tail of a population extending to larger sizes (including objects with HI diameters $> 30'$ not included in this work). We will explore this break in HI size in the HVC population in future work with a larger survey area. The W_{50} values are centered around 15-30 km s^{-1} with a few UCHVCs having widths extending up to 70 km s^{-1} ; the median linewidth is 23 km s^{-1} . There are clouds whose velocities cluster near both $v_{LSR} \pm 120 \text{ km s}^{-1}$, with a much stronger clustering of positive velocity clouds. However, when the MIS UCHVCs are considered, this clustering disappears. The vast majority of negative velocity clouds are also excluded from the MIS UCHVCs; the negative velocity clouds are predominantly in the fall sky, where large scale HI structure is much more prevalent, preventing the inclusion of any UCHVCs into the most-isolated subsample.

3.2.4 Inferred Cloud Parameters

Given the observed properties of the UCHVCs, integrated flux density (S_{21} , Jy km s^{-1}), average angular diameter ($\bar{a} = \sqrt{ab}$, arcminutes) and velocity width (W_{50} , km s^{-1}), it is straightforward to derive some simple properties of the UCHVCs, modulo the unknown distance d (in Mpc), with the assumption that the clouds are optically thin. Sequentially, below we derive the mean atomic density, mean column density, HI mass, indicative dynamical mass within the HI extent, and HI

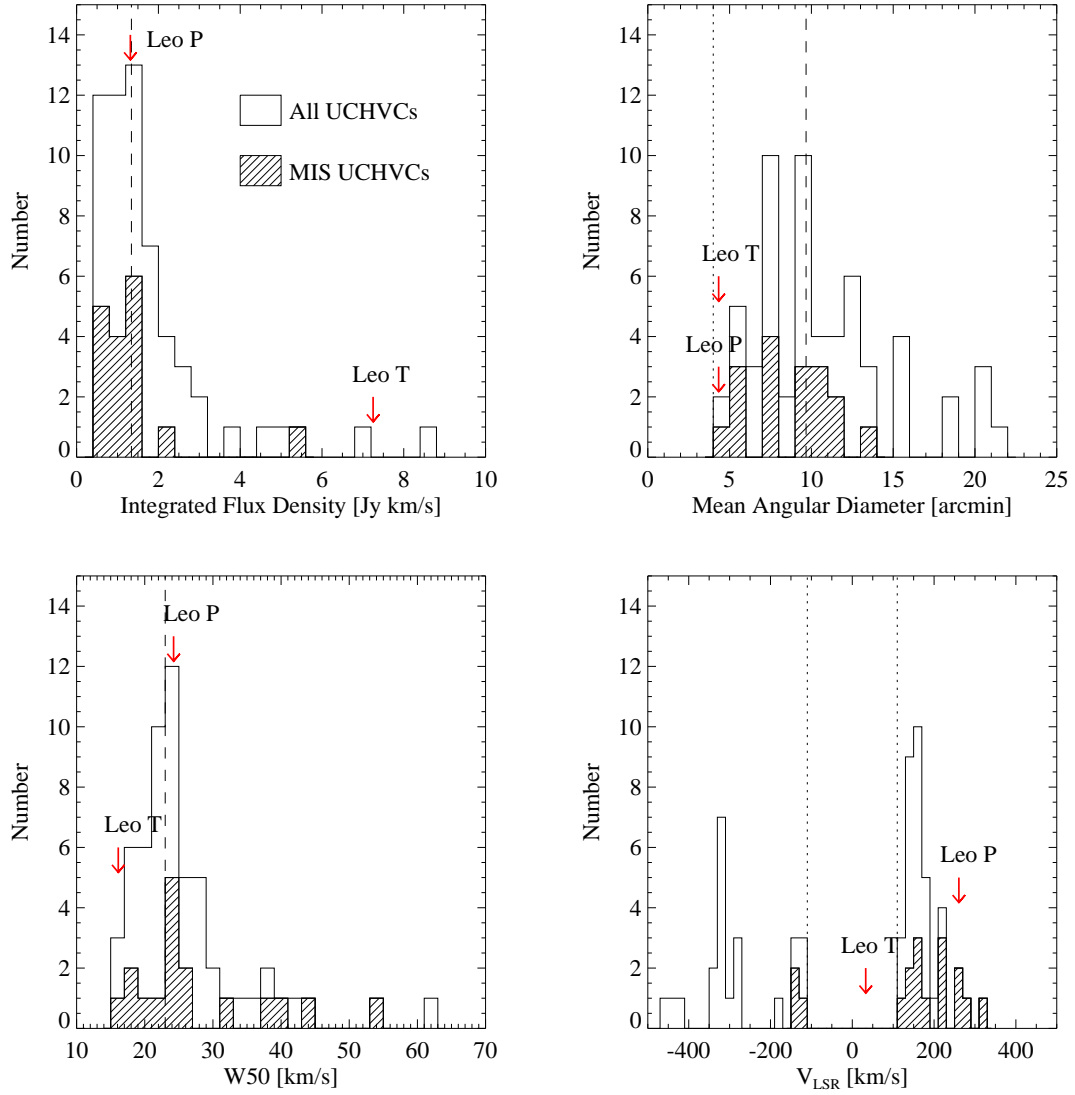


Figure 3.7: Histograms of measured properties for the UCHVCs. Hashed histograms indicate the most-isolated subsample. The measured values for Leo T and Leo P from the ALFALFA data are indicated with red arrows. The dashed lines are the median values of the UCHVCs; the most-isolated subsample has a slightly lower median flux density value and identical median values for the HI size and W_{50} . The dotted lines indicate observational boundaries. In the upper right panel, the dotted line indicates the smallest structure that can be resolved by Arecibo, and in the the bottom right panel the dotted lines indicate the velocity selection criterion.

diameter.

$$\bar{n}_{HI}[\text{atoms cm}^{-2}] = 0.74 S_{21} \bar{a}^{-3} d^{-1} \text{ cm}^{-3} \quad (3.5)$$

$$\bar{N}_{HI}[\text{atoms cm}^{-2}] = 4.4 \times 10^{20} \bar{a}^{-2} S_{21} \text{ cm}^{-2} \quad (3.6)$$

$$M_{HI}[M_{\odot}] = 2.356 \times 10^5 S_{21} d^2 \quad (3.7)$$

$$M_{dyn}[M_{\odot}] = 6.2 \times 10^3 \bar{a} W_{50}^2 d \quad (3.8)$$

$$D_{HI}[\text{kpc}] = 0.29 \bar{a} d \quad (3.9)$$

$$(3.10)$$

Of these derived properties, \bar{N}_{HI} is especially noteworthy as it does not depend on the distance. It should be noted that the column density values derived here are average values based on the global properties of the UCHVCs, in contrast to the approximation of spatially-resolved column density contours in Figure 3.6. Due to the large beam size of Arecibo, these values represent underestimates of the peak values of the clouds. We note that the dynamical mass is an indicative mass dynamical mass only. In addition to the uncertainty in the distance of the UCHVCs, the contribution to the linewidths of the UCHVCs from thermal broadening is unknown. For a range of reasonable temperatures, the thermal broadening can range from 16-21 km s⁻¹. For the clouds with the largest linewidths, the thermal broadening contribution (when accounted for in quadrature) may be negligible, while the narrowest clouds may be fully thermally supported. However, they could still have large-scale motions on the order of the thermal broadening, or less. For example, Leo P has a linewidth of 24 km s⁻¹ and a rotational velocity of 9 km s⁻¹, uncorrected for disk inclination [Giovanelli et al., 2013]. To derive accurate dynamical masses will require higher resolution HI images in which evidence of large scale motions can be discerned (and, of course, distance information).

In Table 3.3, we summarize the inferred properties of the UCHVCs. The columns of the table are as follows:

- Col. 1 and 2: source id as in Table 3.1
- Col 3: HI diameter in kpc at $d = 1$ Mpc (Eqn. 3.9)
- Col 4: log of the mean atomic HI density at $d = 1$ Mpc, in cm^{-3} (Eqn. 3.5)
- Col 5: log of the mean HI column density, in cm^{-2} (Eqn. 3.6)
- Col 6: log of the HI mass at $d = 1$ Mpc, in solar units (Eqn. 3.7)
- Col 7: log of the indicative dynamical mass within D_{HI} at $d = 1$ Mpc, in solar units (Eqn. 3.8)

Table 3.3. Inferred Cloud Properties

Source	AGC	D_{HI} kpc d	n_{HI} $\text{cm}^{-3}d^{-1}$	$\log \bar{N}_{HI}$ cm^{-2}	$\log M_{HI}$ $M_{\odot}d^2$	$\log M_{dyn}$ $M_{\odot}d$
HVC111.65-30.53-124	103417	5.8	-3.68	18.40	5.74	7.74
HVC123.11-33.67-176	102992	4.6	-3.62	18.35	5.48	7.64
HVC123.74-33.47-289	102994	1.6	-2.54	18.98	5.20	7.18
HVC126.85-46.66-310	749141	2.7	-3.12	18.62	5.28	7.48
HVC131.90-46.50-276	114574	2.2	-2.95	18.72	5.22	7.54
HVC137.90-31.73-327	114116	6.2	-3.53	18.58	5.97	8.19
HVC138.39-32.71-320	114117	4.5	-3.07	18.90	6.02	7.67
HVC154.00-29.03-141	122836	6.0	-3.25	18.85	6.21	7.97
HVC205.28+18.70+150★	174540	2.2	-2.46	19.19	5.69	7.40
HVC196.50+24.42+146	174763	3.8	-3.02	18.87	5.82	7.51
HVC196.09+24.74+166	174764	2.2	-2.94	18.71	5.19	7.43
HVC198.48+31.09+165	189054	4.6	-3.49	18.49	5.62	7.82
HVC204.88+44.86+147★	198511	2.0	-2.81	18.81	5.24	6.99
HVC234.33+51.28+143	208315	3.6	-2.72	19.15	6.07	7.49
HVC250.16+57.45+139	219214	1.6	-2.58	18.93	5.12	7.13
HVC252.98+60.17+142	219274	5.8	-3.11	18.97	6.30	7.96
HVC253.04+61.98+148	219276	3.7	-3.14	18.74	5.69	8.01
HVC255.76+61.49+181	219278	2.4	-2.91	18.78	5.33	7.21
HVC256.34+61.37+166	219279	3.2	-3.10	18.72	5.55	7.60
HVC245.26+69.53+217★	215417	2.8	-3.24	18.52	5.22	7.24
HVC277.25+65.14-140★	227977	1.5	-2.64	18.86	5.03	7.24
HVC274.68+74.70-123★	226067	1.3	-2.14	19.29	5.34	7.91
HVC290.19+70.86+204	226165	2.2	-2.83	18.83	5.33	7.32
HVC292.94+70.42+159	229344	4.4	-3.46	18.50	5.59	7.33
HVC295.19+72.63+225	226170	3.8	-3.41	18.48	5.44	7.80

Table 3.3 (cont'd)

Source	AGC	D_{HI} kpc d	n_{HI} $\text{cm}^{-3}d^{-1}$	$\log \bar{N}_{HI}$ cm^{-2}	$\log M_{HI}$ $M_{\odot}d^2$	$\log M_{dyn}$ $M_{\odot}d$
HVC298.95+68.17+270★	227987	3.5	-2.62	19.23	6.12	7.70
HVC324.03+75.51+135	233763	1.8	-2.51	19.05	5.35	7.50
HVC320.95+72.32+185	233830	5.3	-3.69	18.35	5.60	7.78
HVC330.13+73.07+132	233831	1.2	-2.23	19.17	5.17	6.83
HVC326.91+65.25+316★	238713	3.1	-3.12	18.68	5.47	7.65
HVC 28.09+71.86-144★	249393	3.3	-3.25	18.58	5.42	8.11
HVC353.41+61.07+257★	249323	3.2	-3.12	18.69	5.50	7.43
HVC351.17+58.56+214★	249282	1.7	-2.29	19.26	5.53	7.77
HVC352.45+59.06+263★	249283	3.9	-3.46	18.44	5.42	7.92
HVC356.81+58.51+148★	249326	1.6	-2.54	18.99	5.22	7.70
HVC 5.58+52.07+163★	258459	3.0	-3.05	18.74	5.50	7.57
HVC 13.59+54.52+169★	258237	2.0	-2.55	19.08	5.50	7.36
HVC 13.60+54.23+179★	258241	2.9	-3.12	18.65	5.37	7.25
HVC 13.63+53.78+222★	258242	2.1	-2.84	18.79	5.22	7.29
HVC 26.11+45.88+163	257994	2.7	-2.72	19.02	5.68	7.48
HVC 26.01+45.52+161	257956	1.9	-2.40	19.19	5.56	7.41
HVC 29.55+43.88+175	268067	2.2	-2.51	19.15	5.65	7.81
HVC 28.07+43.42+150	268069	2.1	-2.62	19.00	5.43	7.57
HVC 28.47+43.13+177	268070	3.5	-3.22	18.64	5.54	7.48
HVC 28.03+41.54+127	268071	2.7	-2.62	19.13	5.80	8.35
HVC 28.66+40.38+125	268072	3.4	-2.85	19.00	5.87	8.11
HVC 19.13+35.24-123	268213	3.1	-2.77	19.04	5.82	7.28
HVC 27.86+38.25+124★	268074	2.8	-3.00	18.77	5.48	7.51
HVC 84.01-17.95-311	310851	5.4	-3.54	18.51	5.79	7.71
HVC 82.91-20.46-426	310865	2.4	-2.91	18.79	5.37	7.40

Table 3.3 (cont'd)

Source	AGC	D_{HI} kpc d	n_{HI} $\text{cm}^{-3} d^{-1}$	$\log \bar{N}_{HI}$ cm^{-2}	$\log M_{HI}$ $M_{\odot} d^2$	$\log M_{dyn}$ $M_{\odot} d$
HVC 80.69-23.84-334	321318	3.7	-3.27	18.61	5.54	7.62
HVC 86.18-21.32-277	321455	2.8	-2.82	18.93	5.62	7.23
HVC 82.91-25.55-291	321320	2.7	-2.93	18.81	5.49	7.52
HVC 84.61-26.89-330	321351	3.5	-3.37	18.49	5.39	7.52
HVC 92.53-23.02-311	321457	3.8	-3.27	18.63	5.60	7.81
HVC 87.35-39.78-454	334256	2.7	-2.85	18.89	5.57	7.59
HVC 88.15-39.37-445	334257	2.0	-2.80	18.81	5.20	7.31
HVC108.98-31.85-328	333613	2.2	-3.02	18.63	5.11	7.23
HVC109.07-31.59-324	333494	2.7	-2.77	18.97	5.63	7.22

★Part of the extremely isolated MIS subsample

The HI masses, dynamical masses, mean atomic densities and mean column densities of the UCHVCs and the MIS UCHVCs are shown in Figure 3.8. At a distance of 1 Mpc, the HI masses are around $\sim 10^5 - 10^6 M_{\odot}$ and the dynamical masses are $\sim 10^7 - 10^8 M_{\odot}$. This would require the UCHVCs to have an ionized envelope of hydrogen or a substantial amount of dark matter in order to be self-gravitating. As discussed in Section 3.4, these median properties are a good match to the minihalo models of Sternberg et al. [2002]. The median dynamical mass is $10^{7.5} d_{Mpc} M_{\odot}$; this is close to the common mass scale of $\sim 10^7 M_{\odot}$ for the UFDs of Strigari et al. [2008].

3.3 The UCHVCs as a Distinct Population

While the minihalo hypothesis is intriguing for the UCHVCs, we must carefully consider other possible explanations. In this section we examine the possibility of associating the UCHVCs with other cloud populations, including large HVC

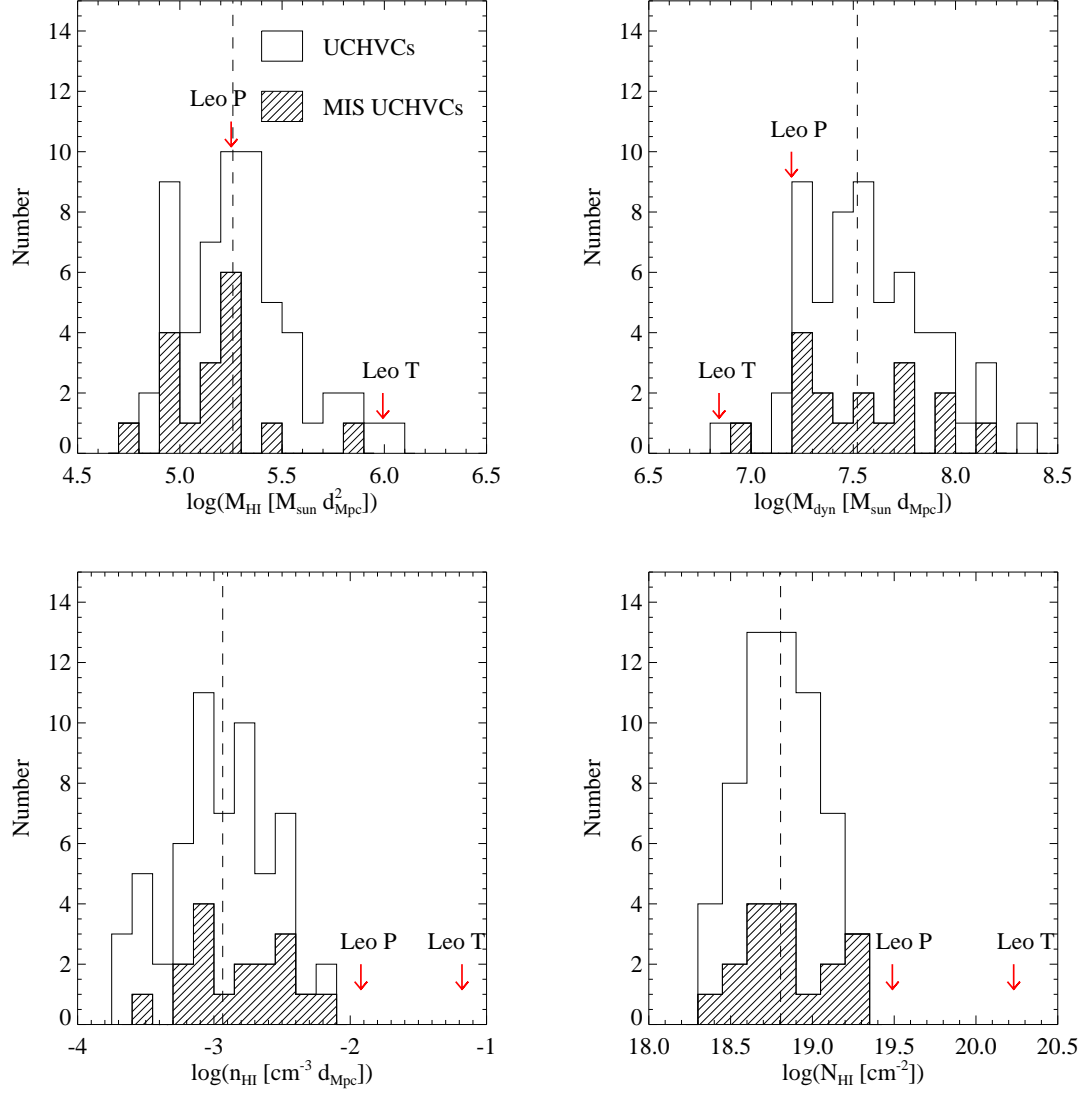


Figure 3.8: The distribution of inferred properties for the UCHVCs. Shading and symbols are the same as in Figure 3.7. The most-isolated subsample has a slightly lower median mass than the full UCHVC sample; for other properties the median values are equivalent between the two samples. Leo T and Leo P are shown for comparison.

complexes, the Magellanic Stream, Galactic halo clouds, and the small cloud populations of the GALFA-HI survey.

3.3.1 The UCHVCs in the Context of Large HVC Complexes

The HVC sky contains many large extended structures composed of multiple clouds. We explicitly require the UCHVCs to be isolated from the known large scale HVC structure of the WvW catalog. However, our isolation criterion for separation from WvW complexes is slightly relaxed in order to avoid excluding potential minihalo candidates. As can be seen in Figure 3.3, the distance to the nearest cloud within a WvW complex can extend to $D = 25^\circ$. As we set our isolation criterion for UCHVCs to a separation of 15° from WvW clouds in complexes, we wish here to consider the possible association of the UCHVCs with WvW complexes. In Table 3.4 we list the UCHVCs that are less than 25° from a WvW complex. We note that only two UCHVCs in the fall sky (HVC86.18-21.32-277 and HVC87.35-39.78-454) are more than 25° from a complex in the WvW catalog; the other fall HVCs not listed in Table 3.4 are separated by less than 25° from clouds associated with the Magellanic Stream in the WvW catalog. Of the 40 spring UCHVCs, seven are potentially associated with known large complexes, the majority of those being with the WA complex. While a few of the UCHVCs may be associated with known large complexes, the vast majority are not, as defined by our isolation criterion.

Table 3.4. UCHVCs within $D = 25^\circ$ of a WvW complex

Complex	UCHVC	Distance to closest cloud degrees
Complex G	HVC111.65-30.53-124	20.1
Complex H	HVC123.11-33.67-176	17.9
Complex ACVHV	HVC137.90-31.73-327	23.8
	HVC138.39-32.71-320	20.9
Complex ACHV	HVC154.00-29.03-141	15.1
Complex WC	HVC205.28+18.70+150	24.6
Complex WA	HVC234.33+51.28+143	16.3
	HVC250.16+57.45+139	19.3
	HVC252.98+60.17+142	21.9
	HVC253.04+61.98+148	24.6
	HVC256.34+61.37+166	24.7
Complex C	HVC 19.13+35.24-123	19.4

Magellanic Stream

The Magellanic Stream (MS) is an extended HI structure first noted by Dieter [1965] and first associated with the Magellanic Clouds by Mathewson et al. [1974]. The MS is generally associated with the disruption of the Magellanic Clouds as they interact with the Milky Way, although the exact mechanisms responsible for the MS are an open area of research. The two main parts of the MS are the Leading Arm (LA), which consists of gas ahead of the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC) in their presumed orbits, and the tail, which consists of the trailing material. Recently, Nidever et al. [2010, hereafter N10] presented an extension of the Magellanic Stream (MS), bringing it to over a 200° length in total. Given the extent of the MS, possible association with the MS must be considered when attempting to understand HVCs of any sort.

For the $\alpha.40$ footprint, the fall sky overlaps the tail of the MS and the spring sky is near the known edge of the LA but not contiguous to it. N10 extended

the known tail of the MS and pointed out its complexity (see their Figure 4), so we must be especially careful with UCHVCs in the fall sky. In Figure 3.9, we show the UCHVCs plotted on the 200° MS presented in N10. The coordinates are the MS-coordinate system of Nidever et al. [2008] based on fitting a great circle to the MS, where L_{MS} is the longitude along the MS and B_{MS} is the latitude above/below the MS. The UCHVCs are shown as large red symbols to increase their visibility; they are not shown to physical scale nor do their colors match the shading of the MS. The top panel shows the HI column density of the MS ($\log N_{HI}$ in cm^{-2}). The bottom panel is the total intensity of the MS integrated along B_{MS} (K deg).

In the spring sky, the $\alpha.40$ footprint approaches but does not overlap the LA of the MS. This lack of direct coverage of the MS makes it a challenge to answer the question: could the UCHVCs be connected to the LA? Future surveys directed at determining any possible continuation of the LA will be able to directly answer this question. Until then, the key to answering this question is determining whether the UCHVCs have compatible velocities to be an extension of the LA. Clearly, the large velocity spread of UCHVCs seen in the bottom panel of Figure 3.9 appears to be incompatible with all of the UCHVCs being associated with the LA. Examining models of the MS can provide insight into these questions. Connors et al. [2006] model the MS as a tidal structure via interaction with the MW and LMC; they predict that the LA extends to $L_{MS} \sim 150^\circ$ with a velocity turn over starting from $L_{MS} \sim 60^\circ$ at $v_{LSR} \sim 300 \text{ km s}^{-1}$ extending to $\sim -150 \text{ km s}^{-1}$. In contrast, Besla et al. [2010] simulate a first passage of the Magellanic Clouds and find a MS that extends to $L_{MS} \sim 50^\circ$ with a velocity increasing with L_{MS} from $v_{LSR} \sim 200$ to 400 km s^{-1} . If the Connors et al. [2006] model correctly represents the history of the MS, then the clouds located at $v_{LSR} < 0 \text{ km s}^{-1}$ could be asso-

ciated with the LA of the MS. If the Besla et al. [2010] model is accurate, then the UCHVCs are generally at higher L_{MS} values than predicted by the model but a few of the positive velocity clouds with $L_{MS} < 100^\circ$ and the highest v_{LSR} values may be associated with the MS. For whichever model of the MS is chosen, some of the UCHVCs could be associated with the LA, but given the large spread in v_{LSR} of the UCHVCs, it is impossible to associate all of the UCHVCs with the LA.

In the fall sky, the $\alpha.40$ footprint overlaps the extension of the MS detailed in N10. In Figure 3.10 we offer a zoomed in view focusing on the fall UCHVCs compared to the MS from N10. Here, there clearly appears to be strong overlap between the UCHVCs and the known MS system. The three clouds in the fall sky at $v_{LSR} > -200 \text{ km s}^{-1}$ appear to be kinematically separated from the MS. Two other clouds at $L_{MS} \sim -100^\circ$ appear to potentially be spatially separated from the MS but the apparent separation could easily be a result of the coverage of observations of the MS. However, it is still possible that some of these UCHVCs do indeed represent galaxies. Many of the UCHVCs that overlap with the MS are also in the direction of the M31 subgroup. Disentangling the gas of known galaxies at a similar velocity from the MS is a long standing problem; see Grcevich and Putman [2009] for illustrative examples. This is also illustrated in Figure 3.10, where several LG galaxies are spatially and kinematically coincident with the MS.

3.3.2 UCHVCs in the Context of Galactic Halo Clouds

Previous studies have uncovered a population of compact clouds associated with the Galactic halo [e.g. Lockman, 2002, Lockman and Pidopryhora, 2005, Stil et al., 2006, Stanimirović et al., 2006, Ford et al., 2010, Dedes and Kalberla, 2010]. While

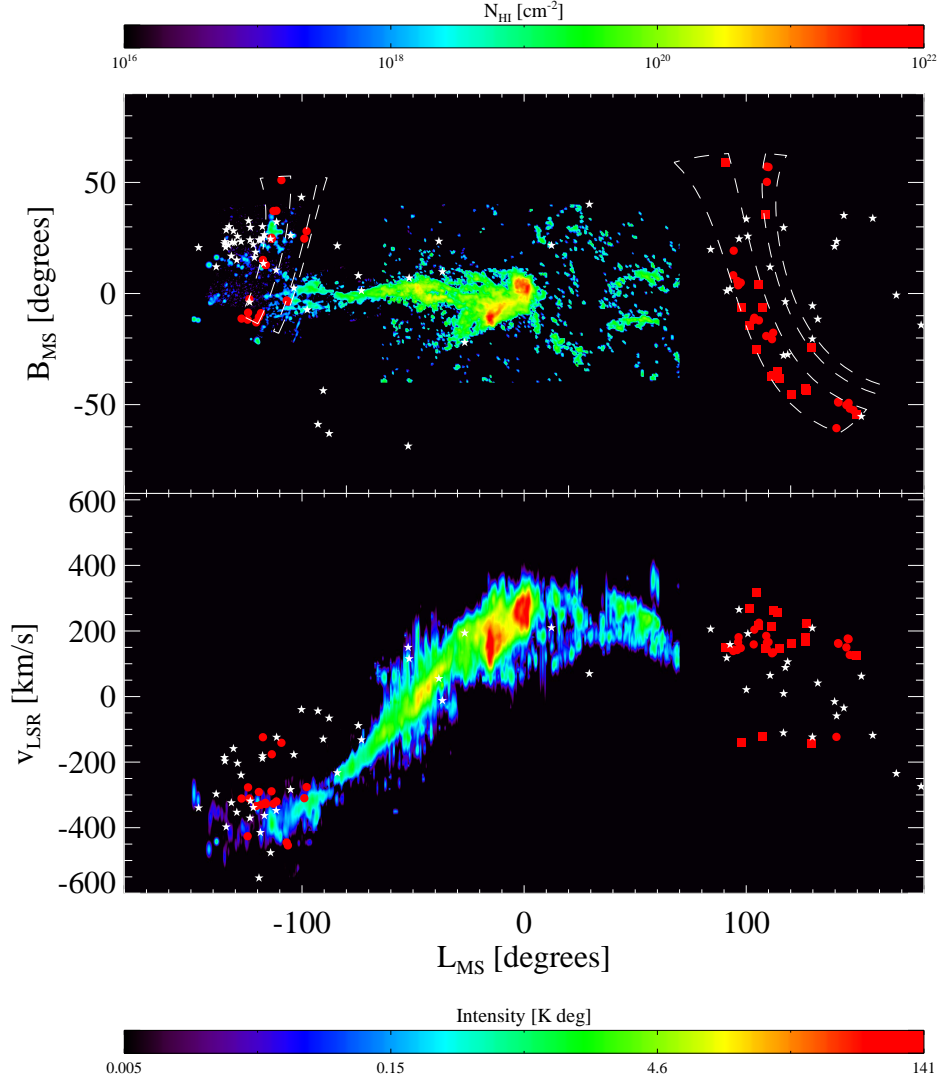


Figure 3.9: The distribution of UCHVCs relative to the MS from N10. Coordinates are those of the MS-centric system from Nidever et al. [2008]. The top panel is the spatial distribution of the MS; the x-axis is L_{MS} and the y-axis B_{MS} . Color coding of the MS indicates the column density, matching N10. The $\alpha.40$ footprint is shown in the top panel by the dashed lines. The red circles are the UCHVCs, the most-isolated subsample is indicated by the squares, and the white stars represent LG galaxies. The bottom panel is the total intensity of the Magellanic HI integrated along B_{MS} (in units of K deg) and shows the kinematics of the MS.

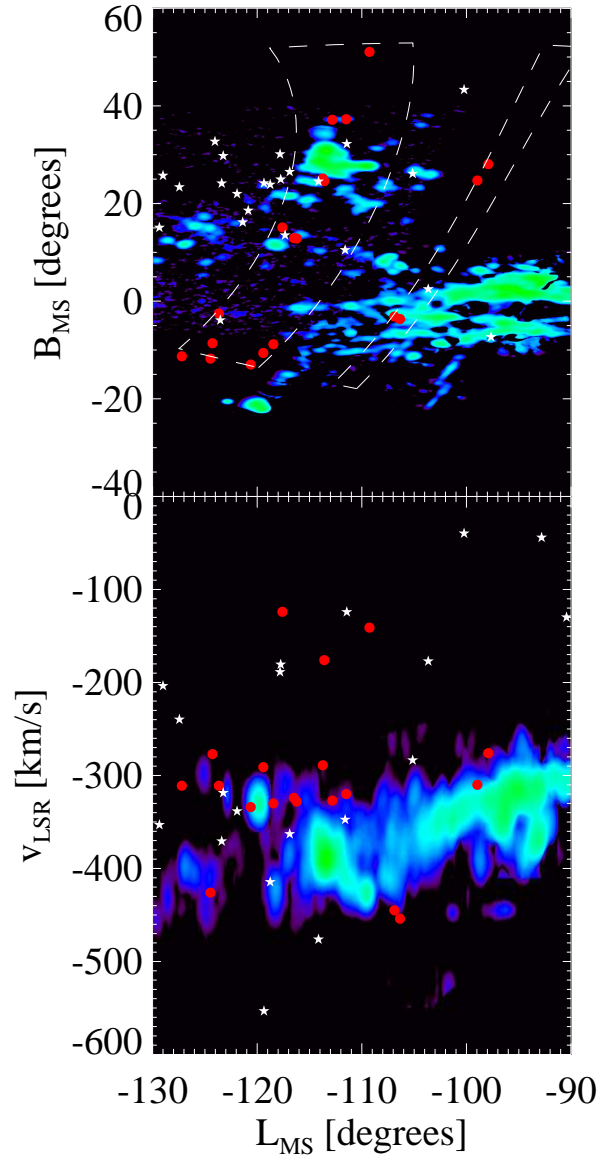


Figure 3.10: A zoomed in view of the fall UCHVCs relative to the MS from Nidever et al. [2010]. Symbols, shading, and panels are the same as in previous figure.

well separated from the Galactic hydrogen, these clouds typically have low v_{LSR} values, and they generally appear to be consistent with Galactic rotation. The Galactic halo clouds with the most extreme velocities of Stil et al. [2006] have v_{LSR} ranging from $\lesssim 100 \text{ km s}^{-1}$ to 165 km s^{-1} . The compact halo clouds also tend to be cold clouds, with the vast majority of reported clouds having $W_{50} < 10 \text{ km s}^{-1}$. Given these characteristics of the halo clouds, the UCHVCs appear as a distinct population. The UCHVCs appear to universally be warm clouds with linewidths greater than 15 km s^{-1} . In addition, many of the UCHVCs have substantial velocities ($|v_{LSR}| > 200 \text{ km s}^{-1}$) that are difficult to account for in a Galactic halo model.

3.3.3 UCHVCs in the Context of the Small Cloud Population of GALFA-HI

GALFA-HI is a survey of neutral hydrogen in the Galaxy which, like ALFALFA, uses the ALFA multi beam receiver on the Arecibo 305m antenna. For GALFA-HI, the IF signal is sent to a different spectrometer than that used by ALFALFA and is restricted to a $\sim 7 \text{ MHz}$ bandpass centered on 1420 MHz . As a result, the GALFA-HI survey has a velocity resolution of 0.184 km s^{-1} and covers a velocity range of $\pm 700 \text{ km s}^{-1}$. It should be noted that much of the GALFA data is taken commensally with the ALFALFA data through the TOGS program. Hence comparison of the results of the two surveys provides a check on our signal processing approach. Begum et al. [2010] presented an initial catalog of compact clouds from the GALFA-HI survey, and Saul et al. [2012, hereafter S12] recently released a catalog of compact clouds for the full initial data release of the GALFA-HI survey.

Herein we focus on the compact clouds of S12 as the most extensive catalog of the compact cloud population discovered in the GALFA-HI survey and examine how the UCHVCs of this work are related.

The initial major differences to note between the catalog of S12 and the UCHVCs are additional selection criteria for the UCHVCs: the limited range of velocities considered and the strong isolation criteria. A vast majority of the compact clouds from S12 do not meet these additional criteria. S12 note several populations of clouds in their catalog which they classify by velocity, linewidth and isolation. They split between warm and cold clouds at a linewidth of 15 km s^{-1} , or a temperature of $\sim 5000 \text{ K}$. It should be noted that while ALFALFA does not have the velocity resolution of the GALFA-HI survey, the velocity resolution of $\sim 10 \text{ km s}^{-1}$ is sufficient to distinguish warm from cold clouds; as can be seen in Figure 3.7, the UCHVCs are all warm clouds with linewidths greater than 15 km s^{-1} . S12 also split their clouds into low velocity and high velocity populations at $|v_{LSR}| = 90 \text{ km s}^{-1}$. They find a few cold clouds with $v_{LSR} > 90 \text{ km s}^{-1}$, but the vast majority of their cold clouds are at lower velocities and associated with the Galactic disk, a very distinct population from the ALFALFA UCHVCs. The populations from S12 of most relevance to this work are their HVC population ($|v_{LSR}| > 90 \text{ km s}^{-1}$) and galaxy candidate population; both of these populations are generally composed of warm clouds. The difference between the HVC population and galaxy candidate population of S12 is that the galaxy candidates have an additional stringent isolation criterion (different from the isolation criteria used here) and hence are the population most directly comparable to the UCHVCs. In Figure 3.11, we compare the distribution of the UCHVCs to the compact clouds of S12 in galactic longitude versus v_{LSR} . In the second Galactic quadrant, the UCHVCs

overlap with the HVCs of S12. This corresponds to the fall sky, and, as noted in the previous section, when considering a stricter isolation criterion for separation from larger HVC complexes akin to that used by S12, the fall UCHVCs cannot be considered isolated structures. In the first and fourth Galactic quadrants, the UCHVCs as a population appear separated from the compact clouds of S12. The positive velocity clouds in the first quadrant and the clouds (at both positive and negative velocities) in the fourth quadrant have no HVC population counterpart in the GALFA compact cloud catalog. Especially in the fourth quadrant, there are multiple clouds at substantial velocities ($v_{LSR} > 200 \text{ km s}^{-1}$) that appear well separated from other clouds populations.

As a check of our methodology and dataset, we also perform a direct comparison of the ALFALFA UCHVCs to the catalog of S12. First, we examine which of the S12 galaxy candidates appear in the $\alpha.40$ catalog. S12 find 28 HVCs that they consider extremely isolated and which they classify as galaxy candidates. Of these, 10 are within the $\alpha.40$ footprint. Two of the GALFA galaxy candidates are classified as extragalactic sources in $\alpha.40$ (AGC191803 and AGC227874) and are clearly associated with optical counterparts; a third S12 galaxy candidate is associated with UGC 7753, a large barred spiral galaxy. Four of their galaxy candidates are within the ALFALFA data but have $|v_{LSR}| < 120 \text{ km s}^{-1}$ and are not included in this work (one is included in the $\alpha.40$ catalog, AGC238801). One of the galaxy candidates is also included here in the UCHVC catalog – HVC351.17+58.56+214. Two of the S12 galaxy candidates are not seen in the ALFALFA data; these are both lower S/N sources ($S/N < 7$) and one is extremely narrow velocity width ($W_{50} = 3.9 \text{ km s}^{-1}$).

Secondly, we can examine the UCHVCs for counterparts in the S12 catalog. 11

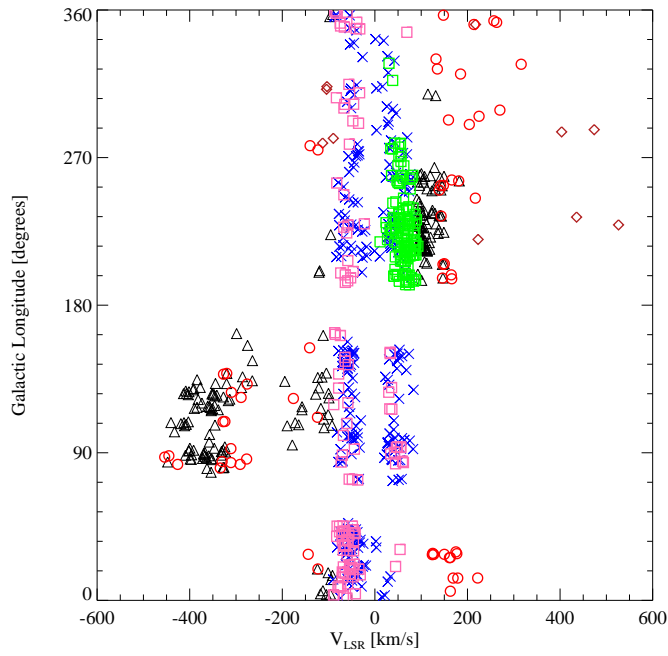


Figure 3.11: The distribution of UCHVCs in $v_{LSR}-l$ space compared to the compact cloud populations of GALFA. Symbols and coloring follow those used in S12: blue Xs are the cold low velocity clouds, pink squares are warm low velocity clouds, green squares are the warm low velocity clouds in the third Galactic quadrant, black triangles are the high velocity clouds, and dark red diamonds are the galaxy candidates. The UCHVCs of this work are shown as bright red circles.

of the 59 UCHVCs are included in the GALFA compact cloud catalog, of which one (HVC351.17+58.56+214) is classified by S12 as a galaxy candidate; the other ten are included in their HVC sample. Seventeen of the UCHVCs are not included in the data coverage of the GALFA DR1 release (D. Saul, private communication); these sources are in the spring sky region of $\delta = 8 - 16^\circ$, where GALFA DR1 has limited coverage because GALFA-HI observations started one year after the commencement of ALFALFA data taking and hence commensal data for that time period are missing. Of the thirty-one UCHVCs with GALFA coverage not contained within the catalog of S12, eight of these sources are found by the algorithm but discarded due to either failing the S12 criteria or data quality issues, such as noise spikes. Five are seen in the data but not found by the signal identification algorithm of S12. The last eighteen are not visible in the GALFA-HI data (D. Saul, private communication 2013). In Figure 3.12, we explore the differences in properties between the UCHVCs found in the dataset of GALFA-HI by the signal identification algorithm of the S12 (including sources discarded from the final catalog), the UCHVCs visible in the GALFA data but not identified by their automated algorithm, and the UCHVCs not visible in the GALFA data. Most strikingly, there is a bimodal distribution in the average column density with the UCHVCs not visible in the GALFA-HI data having the lowest average column densities. In addition, there is a velocity width effect; generally the UCHVCs identified within the GALFA dataset are the narrowest velocity width sources. In the bottom right panel of Figure 3.12, we focus on UCHVCs with integrated flux densities less than 3 Jy km s^{-1} as the higher flux sources are all detected in the GALFA-HI data. Then, there are 18 UCHVCs with linewidths greater than 23 km s^{-1} , the median W_{50} of the full sample. Of these, only three are identified in the GALFA-HI dataset and those still tend to be among the highest flux objects

with integrated flux densities greater than $1.45 \text{ Jy km s}^{-1}$, above the median value of $1.34 \text{ Jy km s}^{-1}$. The UCHVCs that are identified within the GALFA dataset that have flux densities below the median value of the UCHVC sample also have linewidths narrower than the median value of the UCHVCs. This is a straightforward result of the different focus of the two surveys; the GALFA-HI data are designed to detect narrow velocity width HI features associated with Galactic hydrogen while the ALFALFA dataset is designed to detect extragalactic HI sources with wider linewidths. While we will address the completeness and reliability of the UCHVC catalog in future work, we note that six UCHVCs not included in the GALFA catalog have all been confirmed as real HI signals via confirmation observations with the Arecibo L-Band Wide receiver (Adams et al. in prep). In addition, the UCHVCs presented here have strict S/N criteria so the likelihood that many of the UCHVCs are false detections is small. This demonstrates the utility of the ALFALFA dataset, detection algorithm presented here, and the source inspection.

3.4 UCHVCs as Minihalo Candidates

The mismatch between observations of low mass galaxies and simulations of dark matter halos remains an outstanding question in understanding both the cosmological paradigm and galaxy formation and evolution. Is the Λ CDM paradigm incorrect? How does star formation and gas accretion proceed in the lowest mass halos? Finding the lowest mass dark matter halos with baryons can help address these question. In this section, we discuss the possibility that the UCHVCs presented in this paper could represent gas-bearing minihalos. In this context, a minihalo is dark matter halo below the critical mass of $\sim 10^{10} M_{\odot}$ where astrophysical

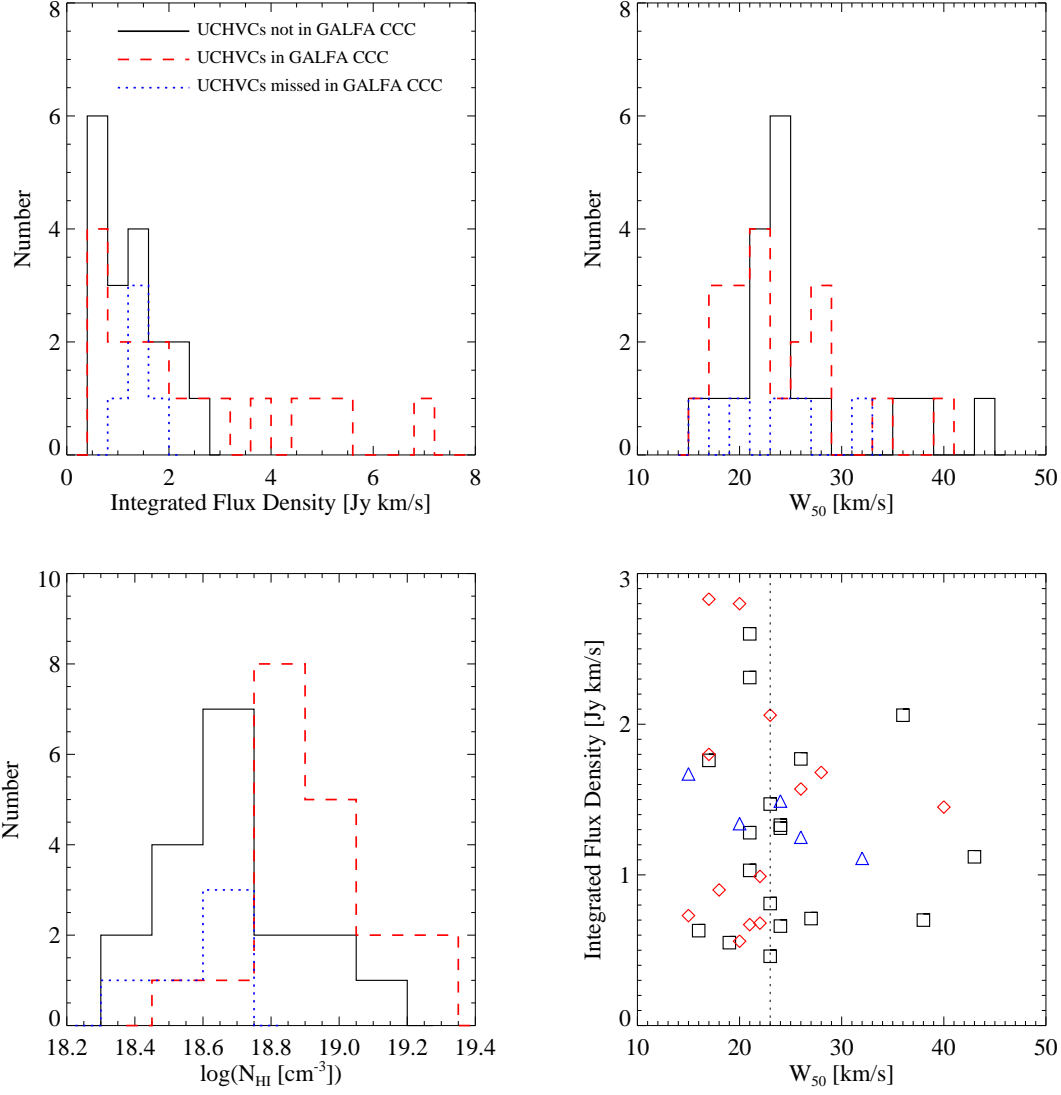


Figure 3.12: Properties for the UCHVCs not seen in the GALFA-HI dataset (solid black lines/squares), UCHVCs included in the GALFA compact cloud catalog (CCC) of S12 and those found by the identification algorithm but discarded from the final catalog (dashed lines/diamonds, red), and UCHVCs seen in the GALFA-HI dataset but missed by the cloud finding algorithm of S12 (dotted lines/triangles, blue). The dotted line in the bottom right panel indicates the median velocity width of the UCHVCs.

processes begin to strongly affect the baryon content [e.g. Hoesft and Gottlöber, 2010, Hoesft et al., 2006]

Sternberg et al. [2002] examined in detail how neutral hydrogen could exist in minihalos. They found that the neutral gas would be surrounded by an envelope of ionized gas, with the specifics depending upon the pressure of the ionized medium the halo is immersed in. They examined both cuspy (NFW) and constant density (Burkert) cores. Cuspy cores are predicted by simulations, while observations of dwarf galaxies indicates that low mass dark matter halos have constant density cores. The UCHVCs appear to match well the Sternberg et al. [2002] minihalo models with a median Burkert density profile, $D_{HI} \simeq 1.4$ kpc, $M_{HI} \simeq 3 \times 10^5 M_{\odot}$, total to neutral gas mass ratio of 15, peak $N_{HI} \simeq 4 \times 10^{19} \text{ cm}^{-2}$, total halo mass $M_{vir} \simeq 3 \times 10^8 M_{\odot}$, surrounded by a hot, ionized IGM of pressure $P_{HIM} = 10 \text{ cm}^{-3} \text{ K}$. The measured column densities are averaged over the size of the cloud and smeared by the $3'.5$ beam of the Arecibo telescope and hence represent a lower limit to the true peak column density, and so they are consistent with the higher peak N_{HI} values of the model. The measured M_{dyn} is an estimate of the total mass within the HI extent; the total size of the dark matter halo exceeds the HI size by a factor of several, explaining the discrepancy between the total halo mass of the model and the inferred dynamical mass from ALFALFA. Work is ongoing to match the individual UCHVC detections to specific individual models (Y. Faerman et al., submitted).

3.4.1 Previous Searches for Minihalos

A LG origin for HVCs, or at least a subset of the HVC population has been considered before. With the advent of large-scale, sensitive, blind HI surveys, interest was revived in HVCs as tracers of dark matter halos. Blitz et al. [1999] and Braun and Burton [1999] both postulated a LG origin for HVCs; Braun and Burton [1999] specifically proposed that compact HVCs (CHVCs), identified by their isolation and undisturbed spatial structure, were good candidates to represent dark matter halos throughout the LG. de Heij et al. [2002a] extracted a set of CHVCs from the Leiden/Dwingeloo Survey (LDS), and Putman et al. [2002] similarly presented a set of CHVCs from the HI Parkes All-Sky Survey (HIPASS). Further work, both observational and theoretical, since the discovery of the CHVC population suggests that they most likely represent a circumgalactic population. The properties of the CHVC population from the two catalogs are summarized in Table 3.5. Sequentially, this table lists: object class, distance (in kpc), HI angular diameter (in arcmin), HI diameter (in kpc), peak column density, W_{50} , integrated flux density, HI mass, and dynamical mass within the HI extent. de Heij et al. [2002b] showed that the properties of the CHVCs for the two datasets are the same when accounting for the better spatial resolution and sensitivity of HIPASS and the better velocity resolution of LDS.

Sternberg et al. [2002] and Maloney and Putman [2003] independently modeled gas in dark matter halos to understand the CHVC population. Based on considerations of their astrophysical properties, both groups concluded that the best interpretation of the CHVCs was as circumgalactic objects at $d \lesssim 200$ kpc. Sternberg et al. [2002] found that if the CHVCs were at $d > 750$ kpc, their dark matter

halos were extremely underconcentrated. They found that at $d \lesssim 150$ kpc, the CHVCs were consistent with being gas pressure confined in dark matter halos. In this scenario, the CHVCs represent the subhalos surrounding the Milky Way from its hierarchical formation. Both pointed out that the gas of the CHVCs must be largely ionized, implying that the total mass of gas is much greater than the observed mass. If the CHVCs were at distances of 0.7-1 Mpc, extremely low dark-matter-to-gas ratios would then be required to match the observed linewidths of the CHVCs, and they would violate the Λ CDM mass-concentration relation. They argued that the CHVCs must be at $d \lesssim 200$ kpc to match size and total dark matter constraints. More recent observational evidence also indicates that the CHVCs must be at circumgalactic distances. The HI masses of the CHVCs at LG distances of ~ 1 Mpc are a few times $10^7 M_\odot$, large enough that they should have been detected in surveys of other galaxy groups but have not [e.g. Pisano et al., 2007, Chynoweth et al., 2011a, Zwaan, 2001, Braun and Burton, 2001, Pisano et al., 2004]. In addition, higher resolution observations of CHVCs show clear ram pressure indicators in many cases, indicating that the CHVCs are located at circumgalactic distances [Westmeier et al., 2005b]. Observations of potential CHVC analogs around M31 also point to a circumgalactic origin. Westmeier et al. [2005a] studied HVCs associated with M31 in high resolution; importantly, the association of these HVCs with M31 allows a distance constraint to be derived. As outlined in Table 3.5, the properties of the M31 HVCs are a good match to the properties of the CHVCs at $d \sim 150$ kpc, indicating that the two samples are likely a similar population.

Multiple searches have been undertaken for minihalos around nearby galaxy groups [e.g. Zwaan, 2001, Braun and Burton, 2001, de Blok et al., 2002, Minchin

Table 3.5. HI Content in the LG - HVCs and Galaxies

Class	d kpc	θ '	D_{HI} kpc	N_{HI} atoms cm ²	W_{50} km s ⁻¹	S_{21} Jy km s ⁻¹	M_{HI} M_{\odot}	M_{tot} M_{\odot}	Refs ^a
UCHVCs	$d = 1000$	10	$2.9d$	$\gtrsim 0.6 \times 10^{19}$	23	1.26	$1.8 \times 10^5 d^2$	$3.3 \times 10^7 d$	1
CHVCs (LDS)	150	60	2.6	1.3×10^{19}	25	102	5.4×10^5	3.5×10^7	2
CHVCs (HIPASS)	150	24	0.52	1.4×10^{19}	35	19.9	1.1×10^5	2.7×10^7	3
M31 HVCs	780	4.6	1.04	3.9×10^{19}	24	2.1	3.0×10^5	4.5×10^7	4
Leo T	420	5	0.6	70×10^{19}	16	6.7	2.8×10^5	$.33 \times 10^7$	5
Leo P	1750	2.0	1.0	20×10^{19}	24	1.31	9.5×10^5	1.3×10^7	6,7,8

^aReferences: 1: this work, 2: de Heij et al. [2002a], 3: Putman et al. [2002], 4: Westmeier et al. [2005a], 5: Ryan-Weber et al. [2008], 6: Giovanelli et al. [2013], 7: Rhode et al. [2013], 8: Skillman et al. [2013]

et al., 2003, Barnes and de Blok, 2004, Pisano et al., 2004, 2007, 2011, Chynoweth et al., 2009, Kovač et al., 2009, Irwin et al., 2009, Chynoweth et al., 2011a,b, Mihos et al., 2012]. Generally, these surveys must choose between sensitivity and coverage area. Irwin et al. [2009] undertook a deep survey of the nearby isolated galaxy NGC 2903 sensitive to an HI mass of $2 \times 10^5 M_\odot$ and covering $150 \text{ kpc} \times 260 \text{ kpc}$. This survey was sensitive enough to (barely) detect a Leo T analog but given that the survey footprint only extends to $\sim 100 \text{ kpc}$ in projected radius from the galaxy center, detection of an object at $\gtrsim 400 \text{ kpc}$ from the galaxy center would depend strongly on orientation. Irwin et al. [2009] did detect one minihalo with an HI mass of $2.6 \times 10^6 M_\odot$, a comparable stellar mass and a dynamical mass of $\gtrsim 10^8 M_\odot$. Chynoweth et al. [2011b] undertook a large ($480 \text{ kpc} \times 1.2 \text{ Mpc}$; $8^\circ.7 \times 21^\circ.3$) survey centered on the region between the M81/M82 and NGC 2403 galaxy groups. Their survey had a mass detection limit of $3.2 \times 10^6 M_\odot$ which is not deep enough to detect a Leo T analog. While their survey covers a large footprint, it is focused on the region between two connected galaxy groups and coverage of the outskirts of the galaxy groups is limited. They detect several massive HI clouds ($M > 10^6 M_\odot$) and determine that these clouds likely arise from tidal processes given their clustering near M81. Mihos et al. [2012] surveyed the M101 group over $1050 \times 825 \text{ kpc}$ ($8^\circ.5 \times 6^\circ.7$) to a mass sensitivity of varying from 2 to $10 \times 10^6 M_\odot$ over their footprint. This footprint includes all objects out to $\sim 400 \text{ kpc}$ from the central galaxy, regardless of orientation, but the survey is not sensitive enough to detect a Leo T analog. They do identify a new low surface brightness dwarf galaxy through an HI detection and a starless HI cloud with an HI mass of $1.2 \times 10^7 M_\odot$.

3.4.2 Known Minihalos in the LG

In considering the UCHVCs as gas-bearing minihalos in the LG, we first want to examine the context of the LG and ask what we may empirically expect a minihalo to look like. The population of the LG has increased substantially in the last few years with the discovery of the UFD satellites of the Milky Way from automated stellar searches of the Sloan Digital Sky Survey [Willman, 2010] and targeted searches for satellites of M31 [e.g. Ibata et al., 2007, McConnachie et al., 2009]. The UFDs have indicative dynamical masses within the baryon extent of $10^6 - 10^7 M_\odot$ and most likely inhabit dark matter halos that qualify them as minihalos. With the exception of Leo T and the recently discovered Leo P, the UFDs are located within the virial radius of the MW or M31 and have no detectable gas content.

Surveys of low mass galaxies in the field indicate that, with large scatter, dwarf galaxies tend to be gas-rich and can have atomic gas as their dominant baryon component [e.g. Geha et al., 2006, Schombert et al., 2001]. Modulo the uncertainties in how astrophysical processes affect the baryon content of the lowest mass halos, one would naively expect the trend of high gas fraction to continue as lower mass galaxies are discovered. Leo T is the only UFD discovered through optical surveys that has neutral gas content; it is also the UFD that is most distant from the MW. The other UFDs are located within the virial radius of the MW or M31 and many show signs of tidal interaction with the MW [e.g. Sand et al., 2012]. Grcevich and Putman [2009] find that morphological segregation is strong in the LG with dwarf galaxies within 270 kpc of the Milky Way or Andromeda showing no evidence of neutral gas content. Leo T is on the edge of detectability for SDSS;

were it located further away, its stellar population would not have been detected [Kravtsov, 2010]. Taken together, these facts raise the possibility that more gas-rich UFDs are lurking in the LG with distances and stellar populations that would leave them undetected in SDSS.

Leo T serves as our prototype of what a gas-rich minihalo will look like; it has motivated our search for more minihalos and the discovery of Leo P. In Figure 3.13 we examine the HI properties of the LG galaxies and neighboring dwarf galaxies within 3 Mpc in comparison to Leo T and Leo P to infer what we may expect for future minihalo detections. The top panel of Figure 3.13 shows a histogram of the HI masses of dwarf galaxies within the LG and neighboring systems, taken from McConnachie [2012]. Leo P and Leo T have some of the lowest HI masses in the LG and Local Volume (LV); we would expect previously undetected systems to have low HI masses. The bottom panel of Figure 3.13 illustrates the parameter space occupied by Leo T and Leo P in the LG and LV; they have low HI masses *and* low dynamical masses.

3.4.3 Evidence for the UCHVCs as Minihalo Candidates

In assessing the UCHVCs as minihalo candidates, we first consider if their astrophysical properties are consistent with the scenario. As mentioned above, the UCHVCs are a good match to the models of Sternberg et al. [2002]. Importantly, the UCHVCs also overcome the objections that ruled out the CHVCs as minihalo candidates throughout the LG. As summarized in Table 3.5, the UCHVCs have HI masses typical of $\sim 10^5 d^2 M_{\odot}$ and HI diameters of $\sim 2.9 d$ kpc. These smaller sizes and lower fluxes suggest that at distances of 1 Mpc, the physical properties

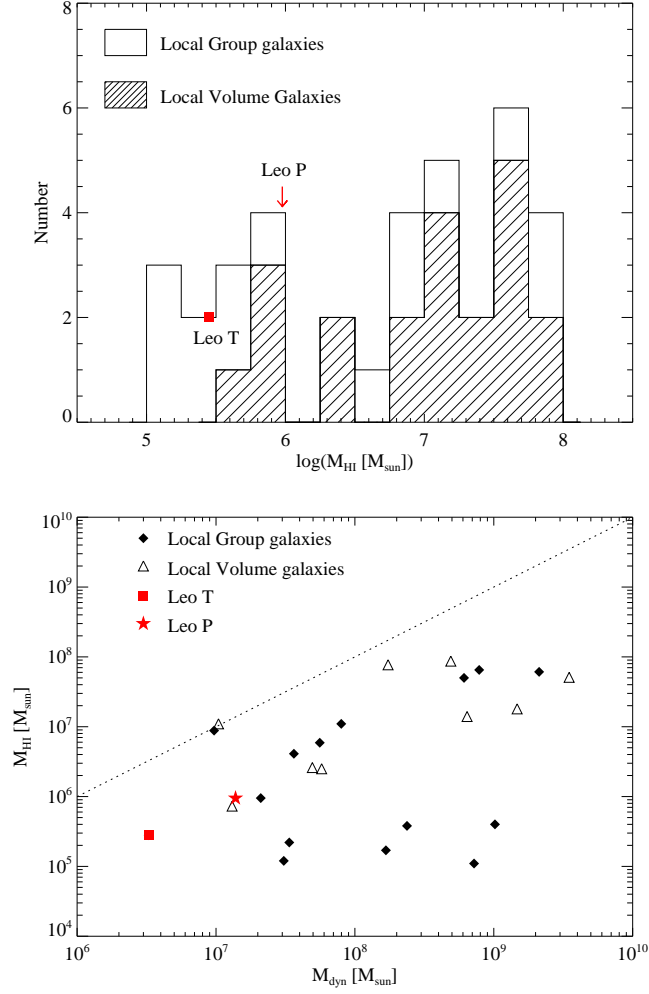


Figure 3.13: The top panel is a histogram of HI mass in the LG and nearby dwarf galaxies in the Local Volume (indicated by the hashed histogram), including Leo T (its contribution indicated by the red filled square), from the catalog of McConnachie [2012]. The location of Leo P is also indicated. The bottom panel is HI mass as a function of dynamical mass within the baryon extent. The diamonds are LG galaxies with HI content, the triangles Local Volume dwarfs, the filled square is Leo T and the filled star is Leo P. The dynamical masses are compiled from the literature and are calculated using a variety of different methods and at different extents of the galaxies; in all cases the dynamical masses are underestimates of the true dynamical mass [Łokas, 2009, De Rijcke et al., 2006, Geha et al., 2010, Shostak and Skillman, 1989, Cook et al., 1999, Hoffman et al., 1996, Mateo, 1998, Ryan-Weber et al., 2008, Kepley et al., 2007, Begum and Chengalur, 2004, Kirby et al., 2012, Skillman et al., 1988, Begum et al., 2005, 2006]. The dotted line indicates where M_{dyn} equals M_{HI} . In addition to having low HI masses, Leo T and Leo P also have low dynamical masses.

of the UCHVCs are good matches to the CHVC properties at distances of ~ 250 kpc. In this scenario, the CHVCs could represent subhalos within the MW and the UCHVCs represent isolated structures within the LG.

The LG is a bound group of galaxies, hence studying the kinematics of the UCHVCs can help constrain their association with the LG. In Figure 3.14 we compare the motions of the UCHVCs to the LG. Following Courteau and van den Bergh [1999], we plot v_{\odot} versus the cosine of the angle from the LG apex. In general the UCHVCs show similar behavior to the motions of the LG galaxies, lending credence to the possibility that they trace LG dark matter halos. They do appear to have a higher velocity dispersion, similar to the nearby neighbor galaxies that are not bound to the LG. This may suggest that the UCHVCs are outlying systems, marginally bound to the LG.

Finally, we offer a preliminary comparison of the UCHVCs to the Via Lactea II (VL) simulation of Diemand et al. [2008], a high resolution cosmological N-body simulation of a Milky Way analog. We compare the spatial and kinematic distribution of the UCHVCs to the dark matter halos of the VL simulation to see if the hypothesis of UCHVCs as minihalos is consistent with theoretical predictions. We utilize the full volume of the simulation, which includes 20,048 halos that extend to more than 3 Mpc from the central MW analog halo. In addition to the central massive halo, there is a second massive halo which is a fortuitous analog to M31 [Teyssier et al., 2012]. In our favored model, we place this second massive halo at the approximate location of M31 in order to most closely match the LG. We also use the original simulation coordinates plus five random orientations of the subhalos to demonstrate the importance of structure within the LG. After the coordinate transformations, we only consider the halos within the simulation that

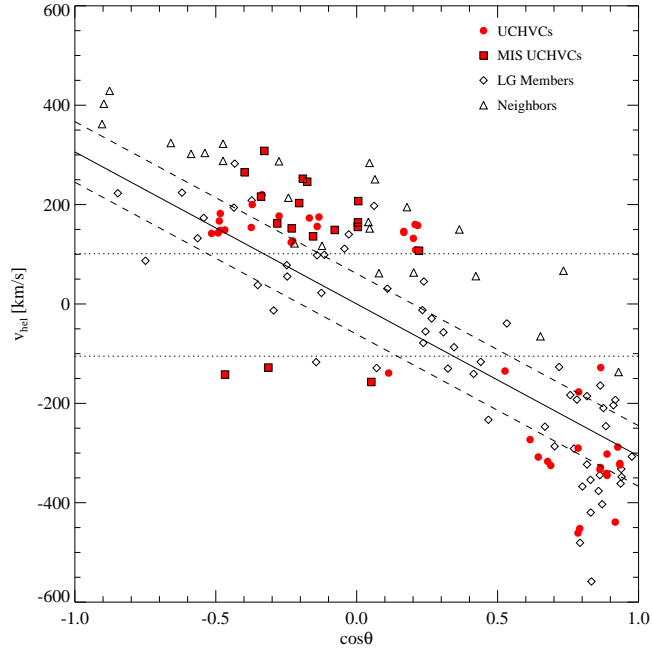


Figure 3.14: Cosine of the angle from the Solar apex versus heliocentric velocity. The solid line shows the relation of Courteau and van den Bergh [1999] and the dashed lines are their stated error. The dotted lines indicate inaccessible velocity space due to the UCHVC velocity selection criterion. The filled red circles are the UCHVCs with the outlined filled red squares indicating the MIS UCHVCs. The diamonds are the LG galaxies from McConnachie [2012] and the triangles are neighboring galaxies within 3 Mpc that are not bound to the LG.

lie within the boundaries of the $\alpha.40$ coverage and meet our velocity criterion.

In Figure 3.15 we show the distribution of galactic latitude and v_{LSR} for the UCHVCs and the VL subhalos. Due to the presence of large and complex HVC structure in the fall sky, we focus on the spring sky for our comparison. In the left column we show all the halos that match our selection criteria; in the right column we show only those halos located further than 250 kpc from the central massive halo to more closely approximate the halos we expect to be gas-bearing. The effects of structure are much more noticeable when only the most distant halos are considered; the different orientations show a much wider spread in the distribution of $|b|$ in this case. The galactic latitude plot is especially important as it provides a quick test of whether the distribution of clouds is within the Galactic disk or a circumgalactic distribution. If the UCHVCs are associated with the Galactic disk, a flattened distribution of $|b|$ values is expected compared to the case if the UCHVCs are distributed around the Galaxy. The UCHVCs and MIS UCHVCs have similar distributions for $|b|$ and $|v_{LSR}|$. The favored orientation of the VL simulation appears to match well the distribution of $|b|$ for the UCHVCs. The large differences in the cumulative distribution function (CDF) of $|b|$ for the random orientations shows the importance of structure. The kinematics of the UCHVCs appear to be consistent with the VL simulation in all cases with the CDF of $|v_{LSR}|$ matching well in all cases. While it is beyond the scope of this paper to do a full halo-population analysis, the rough analysis presented here shows that the UCHVCs agree reasonably well with the VL simulation.

We can also use the VL simulations to provide a rough check of the numbers of halos expected. There are 40 UCHVCs in the spring sky, including 17 in the most-isolated subsample. We compare to our favored orientation of the VL simulation,

noting that it matches the spring sky in that we are looking into the outskirts of the simulation as the spring region of ALFALFA probes the outskirts of the LG. There are a total of 168 VL halos that meet our velocity criterion in the region of the simulation that matches the $\alpha.40$ spring footprint. When limited to halos with distances from the central MW analog halo greater than 250 kpc, there are a total of 44 halos; 27 of these halos have $M_{tidal} > 10^7 M_{\odot}$. Given the roughness of our numbers the two populations appear to be consistent.

3.4.4 The UCHVCs as Galaxies

As galaxies, the UCHVCs would favor the outskirts of the LG, rather than the central regions, with distances of $\sim 500 \text{ kpc} - 1 \text{ Mpc}$. They would have HI masses of $\sim 10^5 M_{\odot}$ with envelopes of warm ionized hydrogen with masses of $\sim 10^6 M_{\odot}$. The indicative dynamical masses within the HI extent are $\sim 10^7 - 10^8 M_{\odot}$, and the total hosting halo masses are likely $\gtrsim 10^9 M_{\odot}$. While this hypothesis is attractive, it cannot be definitively proven until distance constraints are in place for the UCHVCs. Further work is necessary in order to constrain their distances as the ALFALFA HI detection carries no direct distance information. The kinematics of the UCHVCs are dominated by LG interactions, so the velocity cannot offer any insights to the distance. The detection of an optical counterpart can constrain the distance through studies of the stellar population. It is also possible to constrain the distance solely through HI by using synthesis imaging to determine the rotational velocity of the UCHVCs and constrain the distance through the baryonic Tully-Fisher relation [e.g. Giovanelli et al., 2013, McGaugh, 2012]. An alternative to confirming the distance of the UCHVCs directly is to detect UCHVC analogs

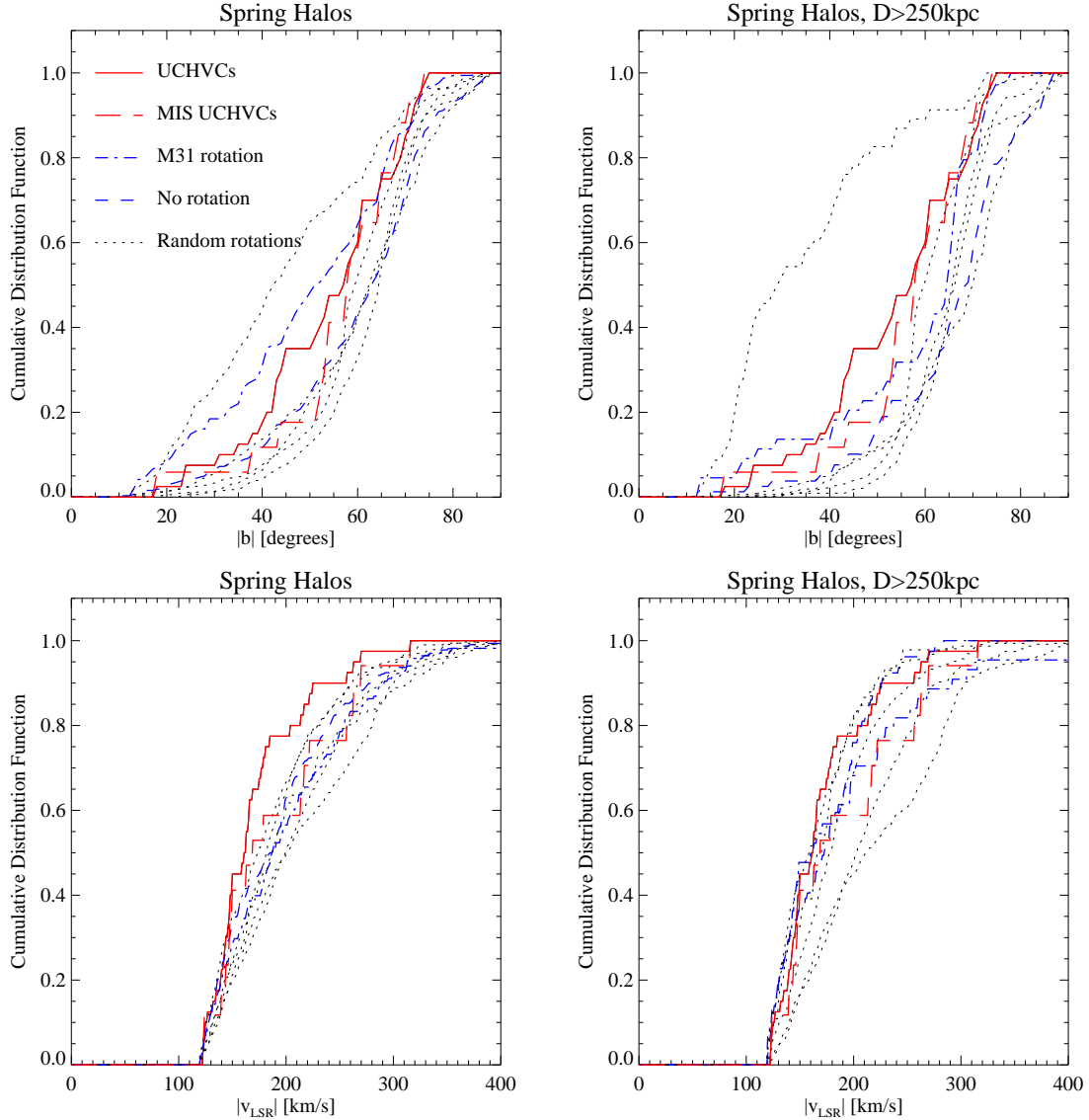


Figure 3.15: The distribution of subhalos from the Via Lactea II simulation compared to the UCHVCs (solid red line) and the most-isolated subsample (dashed red line). The dot-dash blue line represents the subhalos in the original simulation coordinate system; the dashed line (blue) is our favored orientation where the simulation rotated to place the second massive halo at the approximate location of M31. The dotted lines represent five random rotations of the simulation coordinates. The left-hand column shows the distribution of all the VL subhalos in the spring footprint that meet our velocity criterion, and the right-hand column shows the VL subhalos that are located further than 250 kpc from the central massive halo. Overall, the UCHVCs appear consistent with the distribution of halos from the simulation, especially for our favored orientation. Given the large differences between halo distribution depending on the rotation of the simulation coordinates, it is clear that accounting for structure is crucial.

around other nearby galaxy groups and use the association with the group to constrain the distance and properties of the clouds. Planned future HI surveys using phased-array-feeds will be able to robustly detect these objects.

Confirming that a subset of the UCHVCs are galaxies will offer many insights. The UCHVCs will increase the number of low-mass galaxies known in the Local Volume, decreasing the discrepancy between simulations and observations. In addition, the UCHVCs will trace the outskirts of the LG allowing the comparison between simulations and observations to be extended to a larger volume. The UCHVCs will also serve as isolated examples of the lowest mass galaxies, having not yet interacted substantially with the MW. The UCHVCs offer the potential to study star formation in extreme, low metallicity environments as the presence of gas means there is a possibility of star formation. In fact, Leo T has recently formed stars and Leo P has ongoing star formation with one HII region. Abundance measurements of the HII region in Leo P indicate that it is among the lowest metallicity systems known and blind HI surveys may prove to be a promising way to detect low luminosity, extremely metal deficient galaxies [Skillman et al., 2013].

The two confirmed low mass gas-rich galaxies in the Local Volume, Leo T and Leo P, both have high average column densities and small HI angular diameters, as can be seen in Figures 3.7 and 3.8. It may be reasonable to expect then that the most compact and highest column density UCHVCs are the best candidates to represent low-mass gas-rich galaxies. HVC274.68+74.70-123, HVC351.17+58.56+214, and HVC13.59+54.52+169 are in the most-isolated subsample, have average angular diameters $< 7'$, and have $\bar{N}_{HI} > 10^{19} \text{ cm}^{-2}$; we suggest that these are the best galaxy candidates in our sample. One of these candidates, HVC351.17+58.56+214 is also identified by the GALFA-HI survey as a

good galaxy candidate. Notably, it is among the most compact clouds included in this catalog ($7' \times 5'$) and has one of the highest column densities ($\log \bar{N}_{HI} = 19.3$). If we adopt a representative distance of 1 Mpc, it has a HI mass of $3.9 \times 10^5 M_\odot$ and an indicative dynamical mass within the HI extent of $2.1 \times 10^7 M_\odot$.

3.5 Conclusion

We present a set of 59 ultra-compact high velocity clouds which are of interest as speculative minihalo candidates. In brief, the properties of the UCHVCs are summarized below.

- They have HI integrated flux densities from 0.66–8.55 Jy km s^{−1} with a median of 1.34 Jy km s^{−1}, linewidths of 15–70 km s^{−1} with a median of 23 km s^{−1}, and angular diameters of 4–20' with a median of 10'.
- They are selected according to strict isolation criteria. As a result, they are distinct from known HVC populations.
- Their HI sizes and HI fluxes allow them to overcome previous objections leveled against CHVCs as LG minihalos.
- They are consistent with the minihalo models of Sternberg et al. [2002]. At a distance of ~ 1 Mpc, they have HI masses of $10^5 - 10^6 M_\odot$ and dynamical masses within the HI extent of $10^7 - 10^8 M_\odot$. Their total gas masses, including the surrounding ionized envelope, would be $\sim 10^6 - 10^7 M_\odot$ and the total hosting halo masses would be $\lesssim 10^9 M_\odot$.
- As galaxies, they would allow us to probe the outskirts of the LG, study

low mass systems that have remained isolated from the MW, and provide an avenue for indentifying extremely metal deficient galaxies.

CHAPTER 4

HIGH SPATIAL RESOLUTION STUDIES OF THE HI DISTRIBUTION OF ULTRA-COMPACT HIGH VELOCITY CLOUDS

The ALFALFA HI detection of a UCHVC provides the global measurements of the integrated flux density and the linewidth, plus an estimate of the angular diameter. However, no other spatial information is available from the ALFALFA data. As shown in Chapter 2, even as coarse a parameter as the elongation of the HI distribution (i.e., the measured axial ratio of a UCHVC) is not able to be measured accurately from the ALFALFA dataset. Higher spatial resolution observations with a radio interferometer are necessary for the HI distribution to be constrained in any meaningful way. This chapter discusses how spatially resolved HI studies can inform the minihalo hypothesis for the UCHVCs. Imaging and analysis of one early target, AGC268069, acquired with the VLA via shared risk observing are presented as a case study for what interferometric observations can reveal about the UCHVCs. Acquisition and analysis of high resolution HI data is an ongoing project with sources currently being observed (June–August 2013) and another proposal submitted in July 2013 for observations in 2014.

4.1 The Importance of the HI Distribution

Spatially resolved studies of the HI content of the UCHVCs via interferometric observations are a powerful tool for understanding the nature of the UCHVCs. Spatially resolved HI intensity maps can be used to search for signs of interaction with the MW and to constrain models of the hosting dark matter halos. Spatially

resolved velocity fields can uncover evidence of rotation and place stricter constraints on dark matter content.

4.1.1 Evidence for Interaction

If the UCHVCs are at Galactic or circumgalactic distances, they are susceptible to processes such as ram pressure disruption and tidal stripping that will leave signatures in the HI distribution. Observations of the HI distribution of CHVCs were used to help constrain the hypothesis that they represent objects located at distances of $\sim 100\text{--}300$ kpc. Westmeier et al. [2005b] mapped CHVCs with Efstelsberg and found that a majority of their sample showed evidence for head-tail structures, bow-shock shapes and irregular structures. These all indicate that the CHVCs are interacting with an ambient medium and argues for distances of ~ 100 kpc for the CHVC population. Uncovering similar HI morphology in the UCHVCs would indicate that they are also nearby objects at distances of a few hundred kpc. Conversely, if the UCHVCs show a smooth undisturbed HI morphology, that is a strong piece of evidence in favor of the minihalo hypothesis.

4.1.2 Evidence for Rotation

Higher resolution HI observations allow a velocity field to be produced for a UCHVC, revealing the kinematics of the gas as a function of spatial position. An ordered velocity field is evidence of rotation, and a tilted-ring model can be fit to such a velocity field to derive a rotation curve. Gas-rich dwarf galaxies are generally rotationally-supported structures. Specifically, Leo P shows evidence of

rotation on the order of $\sim 11 \text{ km s}^{-1}$ [inclination corrected; Giovanelli et al., 2013]. Leo T shows evidence for a gradient across its velocity field but is a dispersion-supported structure [Ryan-Weber et al., 2008]. The presence of clear rotation in a UCHVC would be considered fairly conclusive evidence that it is dark matter dominated structure.

In addition, the presence of a velocity gradient and/or rotation curve allows the determination of an accurate dynamical mass, modulo uncertainties in the inclination of the disk and distance to the UCHVC. Even for a dispersion-supported structure, higher resolution HI observations allow an improved estimate of the dynamical mass from improved constraints on the size of the HI disk.

4.1.3 Comparison to Models

Observations of the UCHVCs with higher spatial resolution can constrain the radial HI density profile and peak column densities of the UCHVCs. These observations can be used to compare the observed HI in the UCHVCs to models of gas in low mass dark matter halos, constraining the nature of the UCHVCs and their hosting dark matter halos (if they exist). Sternberg et al. [2002] used their models to argue that the CHVCs must be at circumgalactic distances. Recently, Faerman et al. [2013] use a radial HI profile of Leo T to constrain the properties of the hosting dark matter halo.

The presence of multiphase cores with both cold neutral medium (CNM) and warm neutral medium (WNM) is predicted for some gas-bearing low mass dark matter halos in the models of Sternberg et al. [2002]. The size of these cores is much

smaller than the Arecibo beam size and so the spectral signature of the CNM (if it is present) would not be visible in the ALFALFA data or L-band Wide (LBW) confirmation observations (higher spectral resolution). Higher spatial resolution HI observations allow the multiphase core to be spatially isolated and identified. Identifying the presence or lack of a multi-phase core can allow constraints to be placed on models of the UCHVCs.

4.2 A Case Study: AGC268069

Observations of AGC268069 (HVC 28.07+43.42+150) with the *Karl G. Jansky* Very Large Array (VLA) in the C-configuration are presented as a case study for the utility of high resolution HI observations. These data were taken as part of a shared-risk observing program during commissioning of the new VLA correlator, WIDAR. This source is among the high \bar{N}_{HI} and compact groups of sources discussed in Chapter 5. Specifically, it has $cz = 132 \text{ km s}^{-1}$, $W_{50} = 29 \text{ km s}^{-1}$, $\theta_{HI} = 7'$, and $S_{int} = 1.15 \text{ Jy km s}^{-1}$. Figure 4.1 shows the ALFALFA HI contours of this source along with the ALFALFA spectrum.

In addition, there are observations with the *Cosmic Origins Spectrograph* (COS) on the *Hubble Space Telescope* (HST) of a background quasar located $10'$ from the source center. The quasar lies outside the extent of the observed HI but may provide a line of sight through an extended envelope of ionized hydrogen (HII). The central HI (of any UCHVC) is presumably surrounded by an ionized envelope that shields the inner material from the metagalactic UV radiation field, allowing it to remain neutral [Sternberg et al., 2002]. Comparisons of different ion-

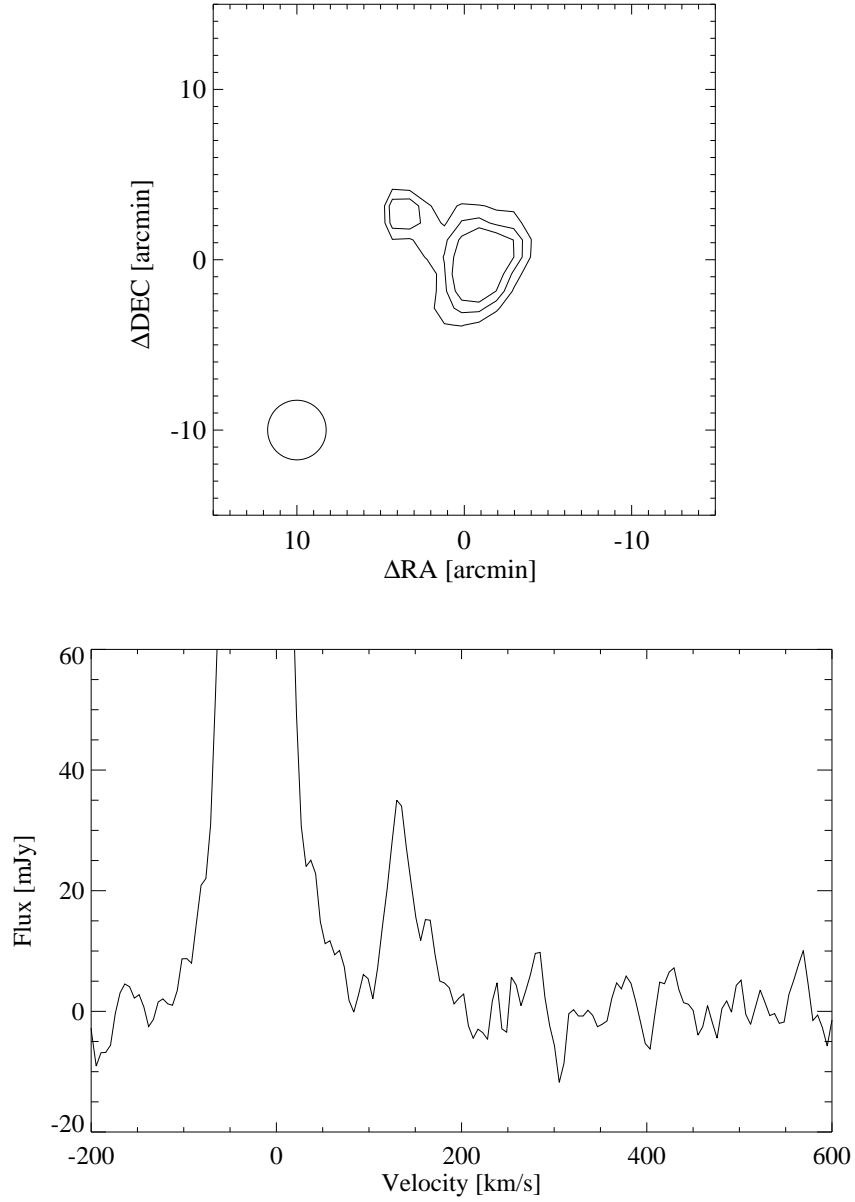


Figure 4.1: *Upper Panel:* The HI contours of AGC268069 from the ALFALFA data at levels of $[5, 7.5, 10] \times 10^{18} \text{ atoms cm}^{-2}$. The field of view is $30'$ and the beam is shown in the lower right corner. *Lower Panel:* The ALFALFA spectrum of AGC268069; it is clearly distinct from Galactic HI emission.

ization levels of atoms within the interstellar medium (ISM) of the UCHVC can allow strict constraints on the hosting dark matter halo and bounding pressure of the surrounding medium. There was no detection of absorption lines from AGC268069 in the spectrum of the background quasar, indicating that the outer envelope of ionized material is truncated at a smaller radius. This implies that the environment of AGC268069 has a bounding pressure greater than $P_{HIM} \sim 10 \text{ cm}^3 \text{ K}$ [Gnat et al., 2013]. The VLA observations presented here are designed to constrain the HI distribution and peak column density to allow more accurate model-based results to be derived from the HST-COS observations.

4.3 The Observations

A total of 26 hours was awarded for this study. The observations of AGC268069 were broken into five different blocks in order to accommodate both the amount of time for which AGC268069 was visible and the dynamic scheduling of the VLA. These five blocks ranged from 4.5-6.5 hours in length; see Table 4.1 for a brief summary of the observing blocks. In all cases, a flux calibrator (J1331+303/3C286) was observed for ~ 15 minutes at the start of the observing block and ~ 7 minutes at the end of the observing block to constrain the flux scale and the bandpass response of the observations. The source is observed in $\sim 20 - 21$ minute blocks sandwiched by ~ 3 minute observations of a phase calibrator (J1553+1256) used to calibrate the phase of the source.

Table 4.1. Summary of VLA/C Observations of AGC268069

Block	Date	Block Time	On-source Time
Block 1	2012 Mar 13	6.5 hrs	~5 hrs
Block 2	2012 Mar 22	5 hrs	~3.75 hrs
Block 3	2012 Apr 3	5 hrs	~3.75 hrs
Block 4	2012 Apr 7	5 hrs	~3.75 hrs
Block 5	2012 Apr 18	4.5 hrs	~3.5 hrs

Table 4.2. Observational Set-up

Parameter	Value
Spectral windows	16
Spectral window bandwidth	250 kHz
Channel width	1.953125 kHz (0.41 km s ⁻¹)
Polarizations	RR, LL
Total bandpass	~2 MHz (~400 km s ⁻¹)

4.4 Data Preparation

This section provides a brief description of the calibration of the VLA data; a complete reduction log of a single scheduling block (including CASA commands) is available in Appendix B. Each scheduling block of data is calibrated individually and the scheduling blocks are combined in the imaging stage (Section 4.5). All calibration was carried out within CASA¹.

¹Common Astronomy Software Application

4.4.1 Calibration

The first step of calibration is to flag the data. Using the `viewer` in CASA, the data is first examined for bad baselines and antennas by looking at amplitudes as a function of baseline pair (sorted by antenna) and time for a single channel. Any bad antennas or baselines are flagged. In addition, amplitudes are examined as function of channel and time for a set of baselines to quantify the presence of any RFI. For these data, the RFI is minimal², and the flagging of any RFI occurs in the next stage with the use of `plotms`. First, amplitudes are plotted as function of time in `plotms` to identify bad times. Then phases are plotted as function of time (for the calibrators only) to ensure that all antennas are properly recording phase. Finally, the amplitudes are examined as a function of frequency to allow RFI to be identified and flagged.

After flagging the data is split by polarization before calibration. This is because in dual-polarization datasets, CASA requires both polarizations to be present to calibrate the antenna. Often, an antenna will be flagged in a single polarization but not the second so splitting the dataset allows as much data as possible to be retained. The flux scale is then set by using a stored model for the flux calibrator. The data is bandpass corrected and then the complex gains (phases and amplitudes) are calibrated. The flux scale is then transferred from the flux to the phase calibrator. Finally, the calibration is applied to the data. The calibrated data is examined in `plotms` to ensure that the calibration looks reasonable. Typically this is done by plotting the calibrated phases against the calibrated amplitudes for the flux and phase calibrators; the calibrated data should cluster in a circle as the calibration sources are point sources. If necessary, further data is

²As expected since the 21cm line is a protected frequency band

flagged and the calibration is redone. The final corrected source data is split into its own file. Continuum sources are then subtracted from the final calibrated uv data. The two separate polarizations are recombined in the imaging stage.

4.4.2 Dealing with Overlapping Passbands

Before imaging the data it is necessary to determine how to deal with the overlapping passbands of the data that are a result of our correlator setup. These data were taken during commissioning of the VLA in a shared-risk mode, meaning the correlator had limited setups available. In order to cover our full frequency range at the desired spectral resolution, we had to use multiple frequency windows in conjunction with each other. A total of 8 spectral windows (SPWs) of 250 kHz each were used to cover the full frequency range of 2 MHz (see Table 4.2). In order to account for the decrease in sensitivity due to bandpass shape at the edge of the individual SPWs, a second set of SPWs were observed offset in frequency by half the frequency width of a single SPW. This ensured good coverage at all frequencies. Figure 4.2 shows the bandpasses of the different SPWs for a single antenna demonstrating the overlapping coverage.

Given the overlapping SPWs of our dataset, we must determine the best way to combine the different SPWs to produce the best data cube. There are three ways that one could think of combining the data:

- Average all channels together
- Average all the channels with good throughput together
- “Glue” the different SPWs together to create uniform frequency coverage

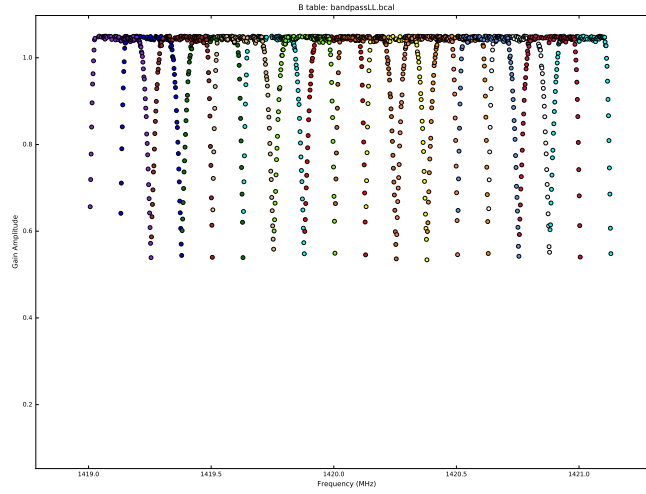


Figure 4.2: The bandpass response of all the spectral windows as function of frequency. Each SPW is colored separately. The overlapping nature of the SPWs is clearly visible.

In the first method, we are assuming that the increase in noise due to fall-off in the bandpass shape is small and that simply combining all data will produce a good data cube. In the second method we are avoiding the noisy edge channels and keeping all other data. In the third method we are attempting to create the most uniform data cube possible. For the third method, it should be noted that we cannot do a simple combination of different SPWs and channels as our channels are narrow enough that signal moves from one channel to another during the course of our observations as a result of the Earth’s motion. Instead, the creation of a single bandpass for the data is done with `cvel` which accounts for the shift in signal due to the Earth’s motion.

Figure 4.3 shows the noise for a single block of observations as function of frequency. Examining the figure, it is clear that the most uniform noise properties come by “gluing” the frequency coverage together with `cvel`. For the remain-

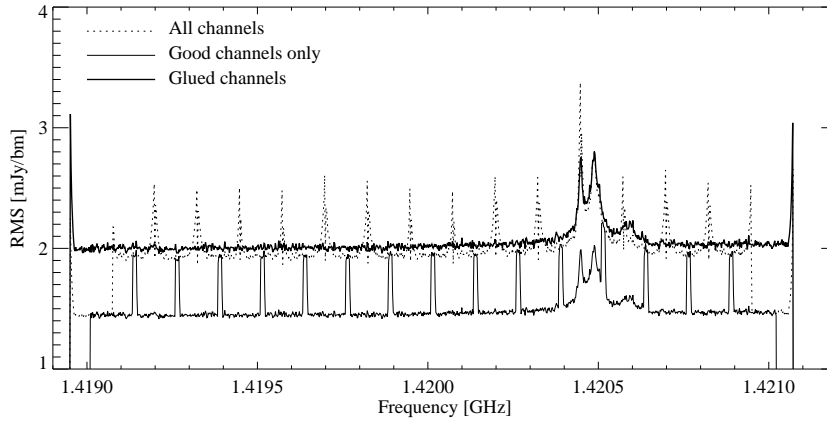


Figure 4.3: The noise as a function of frequency for the three different methods for combining the spectral window (SPW) frequency coverage. “Gluing” the frequency coverage together produces the most uniform data cubes, although it does come at a cost of slightly higher noise.

der of this work, data cubes are created from visibilities that have been “glued” together.

4.5 Imaging

Radio interferometer arrays sample the spacing of a single radio dish of the size of the largest baseline separation between two antennas. Fundamentally, the data are sampled points in the uv plane of the Fourier transform (FT) of the true sky image of the source, and there is no unique way to create an image of the source. The most common way to create an image from interferometric data is “cleaning”. A “dirty” cube is created by taking the FT of the measured uv data, typically via the fast Fourier transform (FFT). This results in a cube that contains the source convolved with the dirty beam of the telescope (the FT of the uv sampling). To create a clean cube, a clean component list (initialized to empty) and a residual

map (initialized to the dirty cube) are created. Then a loop process is started until a threshold or set number of iterations is reached. The loop is as follows:

- The strongest feature in the residual map is identified as a point source.
- A fraction of this source (gain, typically 0.2) is added to the clean component list.
- The fraction of the source is subtracted from the residual map.
- Identify the strongest feature in the new residual map and continue.

When the stopping point is reached, the clean component list is convolved with an estimate of the main lobe of the dirty beam and added to the residual map to create the “clean” image. Miraculously this methodology works extremely well, even for extended low surface brightness sources, although alternative methods have been developed that are specifically optimized for extended emission.

One such alternative is `msclean` (multi-scale clean) which operates similarly to `clean` but fundamentally assumes that sources in the sky are extended structures of different spatial scales [Cornwell, 2008]. `msclean` operates more efficiently than `clean` by reducing the number of iterations needed, both by allowing extended sources to be modeled more efficiently with spatially extended functions and because the gain for subtracting a source can be set to a higher value without divergence. It has been demonstrated that `msclean` returns images with few artifacts [Rich et al., 2008]. For these reasons, we will use `msclean` to produce clean images in this work.

4.5.1 Robustness and Tapering

One important parameter to consider when imaging interferometric data is how the uv data is gridded and weighted for the FFT. It is possible to weight the data and baseline spacings in different manners which can result in significantly different resolutions as the synthesized beam of the telescope (the FT of the uv sampling) changes. In this section we explore two different parameters that affect the resolution and other properties of the final data cubes: robustness and tapering.

The two extremes for weighting uv data are natural and uniform weighting. Natural weighting gives more weight to short baselines and maximizes point source sensitivity while degrading resolution. Uniform weighting gives more weight to longer baselines and fills the uv plane more evenly, resulting in lower outer sidelobes and higher resolution but it also decreases the point source sensitivity. The “robustness” parameter was originally proposed by Briggs and introduces a range of weighting schemes ranging between the two extremes. In CASA, the robustness parameter can range between $[-2, 2]$ where -2 is close to uniform weighting and 2 is close to natural weighting. A robustness of 0 is typically a happy medium in the trade-off between sensitivity and resolution.

In order to choose the best set of robustness parameters to use, a single block of observations was imaged with the full range of robustness values from $[-2, 2]$ in steps of 0.1 , plus natural and uniform weighting. Figure 4.4 shows the beam size and rms (averaged over the channels imaged) as function of robustness. The natural weighting has the lowest RMS of 1.45 mJy/beam which is important for the low surface brightness emission of the UCHVCs. In addition, at a robustness of 0.2 the average beam size has decreased by 25% to $13''$ from $17''$ while the noise

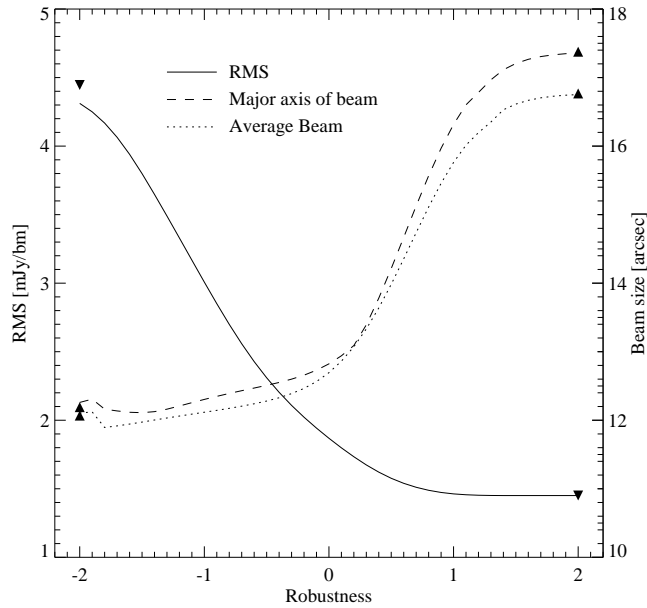


Figure 4.4: The RMS (solid line), major beam axis (dashed line) and average beam size (dotted line) as a function of different robustness values for a single observing block. RMS values for natural and uniform weighting are indicated by downward triangles, and upward triangles indicate beam sizes for the two weighting schemes.

has only increased by 20% to 1.73 mJy/beam. Subsequent imaging and analysis will include two cubes, one produced with natural weighting and the other with a robustness value of 0.2.

In addition to weighting the grid of uv data, a taper can be applied. A taper multiplies the gridded data by a 2D gaussian with specified dimensions, increasing the effect of the shortest baselines (longest spatial scales). This results in a much larger beam which is much more sensitive to extended emission. This is akin to smoothing in the image plane (convolution by a Gaussian) but happens in the uv plane during the imaging process. We image a third data cube with natural weighting and a symmetric taper of $6\text{ k}\lambda$.

Table 4.3. Properties of VLA Data Cubes with Different Imaging Schemes

Weighting	Taper $k\lambda$	Beam " \times "	δv km s^{-1}	RMS ^a mJy/bm	Scale ^b pixels	F_{int}^c Jy km s^{-1}	F_{int}^d Jy km s^{-1}
Rob=0.2	—	14.3×13.5	3.34	0.84	[0,3,12]	0.63	0.39
Natural	—	18.8×16.2	3.34	0.71	[0,4,16]	0.52	0.41
Natural	6	26.5×25.3	3.34	0.85	[0,5,20]	0.52	0.40

^aRMS per channel in the dirty cube.

^bScales used for multiscale clean in pixels; in all cases pixels are 5".

^cIntegrated flux density for the velocity range [113.7, 153.8] km s^{-1} .

^dIntegrated flux density for the velocity range [120.4, 143.7] km s^{-1} .

4.5.2 Imaging AGC268069

AGC268069 is imaged in the three different weighting schemes discussed over the frequency range 1419.5 – 1420.0 MHz (87 – 190.5 km s^{-1}) with three different frequency resolution: 7.9125 kHz, 15.825 kHz, and 31.65 kHz (1.67, 3.34 and 6.68 km s^{-1}). The pixel size is 5" and images are 512×512 pixels. All cubes are cleaned to the 1 sigma level using a multiscale clean; the scales used depend on the beam size for that imaging scheme and are listed in Table 4.3 in pixel units. Table 4.3 also lists the final beam sizes for the three differently imaged data cubes.

After creating all the image cubes, a universal mask to isolate the source is created from the tapered data cube with a resolution of 15.825 kHz. Full details of the masking procedure are given in Appendix B.2.2.

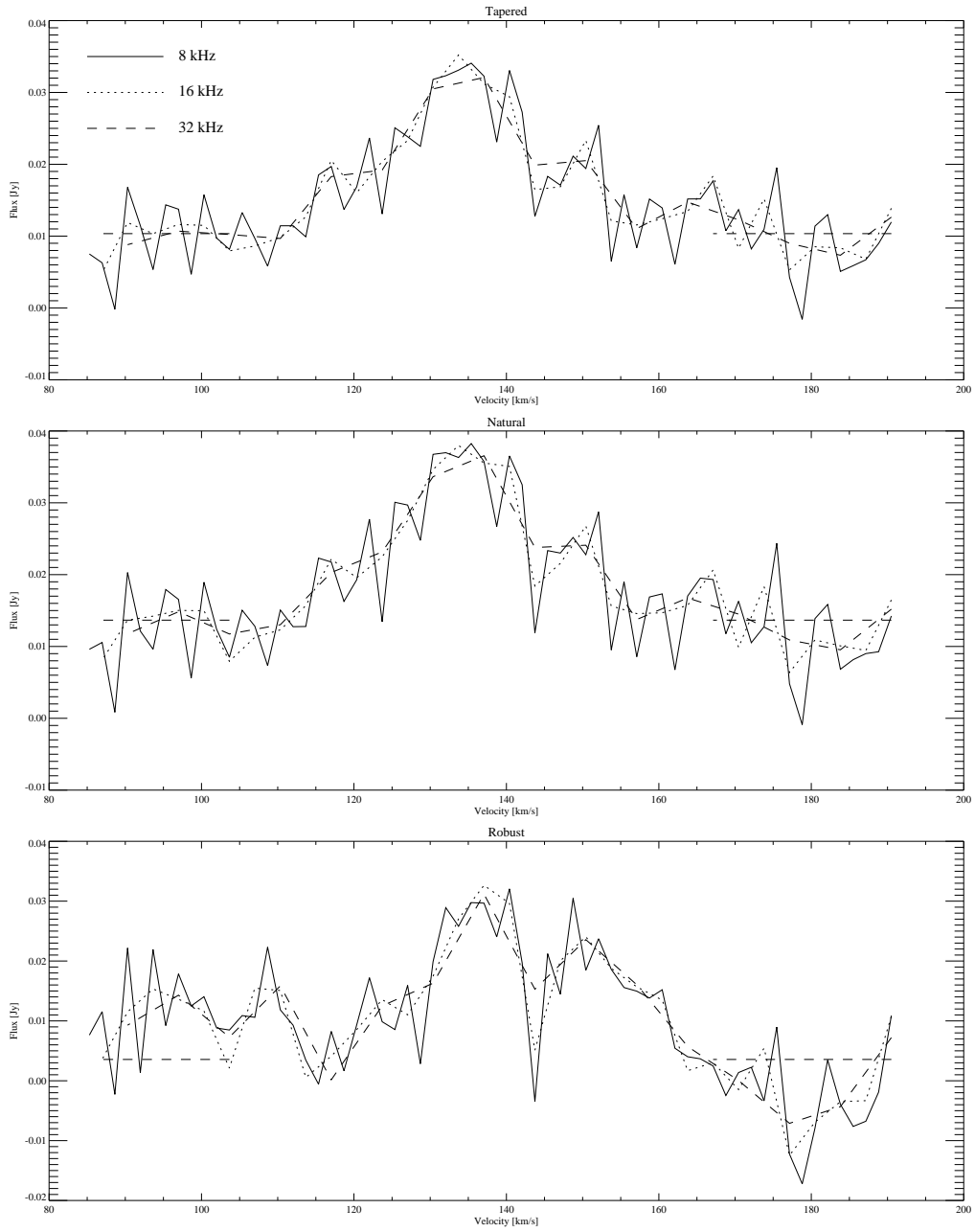


Figure 4.5: The spectra for all three imaging schemes and all three velocity resolutions. Horizontal dashed lines indicated the fitted pedestal flux level and the range of velocities used to fit for it.

Spectra of AGC268069 and Evidence of Pedestaling

The mask was applied to all nine data cubes, and Figure 4.5 shows the resulting spectra. The 16 kHz and 32 kHz spectra show very similar behavior while the 8kHz spectra are much noisier. It is also evident that the baseline of the spectra are elevated, and the amount of elevation depends upon the imaging scheme used to produce the data cube. This is most likely the result of deep cleaning resulting in an artificially elevated (“pedestal”-ed) flux level. In order to compensate for this, the regions of the spectra with no signal ($[87.0, 103.7]$ and $[167, 190.5]$ km s^{-1}) were used to fit the pedestaled flux level; full details are available in Appendix B.2.3. As the 16kHz cubes appear to be the best trade-off between sensitivity and spectral resolution, they are the only cubes for which the pedestal effect is removed from, and they are the cubes that will be used in the rest of this chapter. Figure 4.6 shows the spectra extracted from data cubes that have had the pedestal effect removed.

Creation of Integrated Flux Density Maps

Two different velocity ranges are used to create moment zero (integrated flux density) maps: $[113.7, 153.8]$ km s^{-1} and $[120.4, 143.7]$ km s^{-1} . The larger velocity range is obtained by examining the spectra in Figure 4.6 to determine where there appears to be emission. The smaller velocity range is the range of velocities where inspection indicates there is real signal that is contiguous across channels; these are the channels that were used to help create the mask (Appendix B.2.2). Both velocity ranges can be seen in Figure 4.6. Figure 4.7 shows the moment zero maps for the (pedestal corrected) 16kHz cubes for both velocity ranges. Generally all maps show the same structure (discussed in detail in §4.6.1) with the restricted velocity

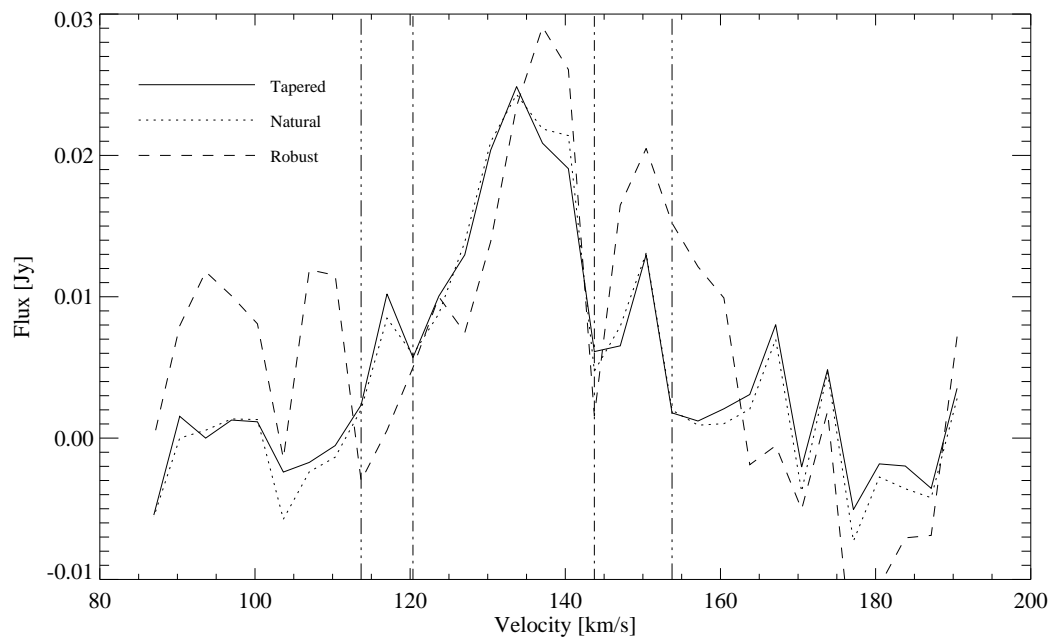


Figure 4.6: The spectra for the 16kHz data cubes after removing the pedestal level. The outermost set of vertical lines indicate the furthest spectral extent to which there appears to be signal; the innermost set of vertical lines indicate the frequency range over which inspection (during the creation of the mask) indicates that there is real signal contiguous across channels.

maps having less intense peaks of emission, both positive and negative. There appears to be a main component composed of three bright cores plus a fainter extended emission to the north. The robustly-weighted map reveals the highest column density peaks (as is expected since it has the smallest beam) of $\sim 6 \times 10^{20}$ atoms cm^{-2} over the full velocity range and $\sim 4 \times 10^{20}$ over the restricted velocity range. The tapered cube produces the cleanest map that clearly shows the presence of three cores of HI plus a northern extension of emission; this cube will be used as the focus for discussion in Section 4.6.

4.5.3 Flux Recovery

A well-known problem with interferometric data is the difficulty of recovering all the flux observed in single-dish observations. Objects with extended emission (like the UCHVCs) tend to have flux “resolved out” by interferometers. First we offer some brief commentary on the single-dish flux value(s) for this source and then detail the flux recovery from the VLA observations presented here.

Single Dish Flux

Generally, the ALFALFA reported fluxes of extragalactic HI detections are consistent with pointed single-dish observations, indicating that the ALFALFA fluxes are accurate [Haynes et al., 2011]. However, the UCHVCs are extended sources and the measured flux is sensitive to the size of the box used during source measurement. AGC268069 is especially sensitive to the source box used for flux measurement as it is located in a region with only a single pass coverage and hence

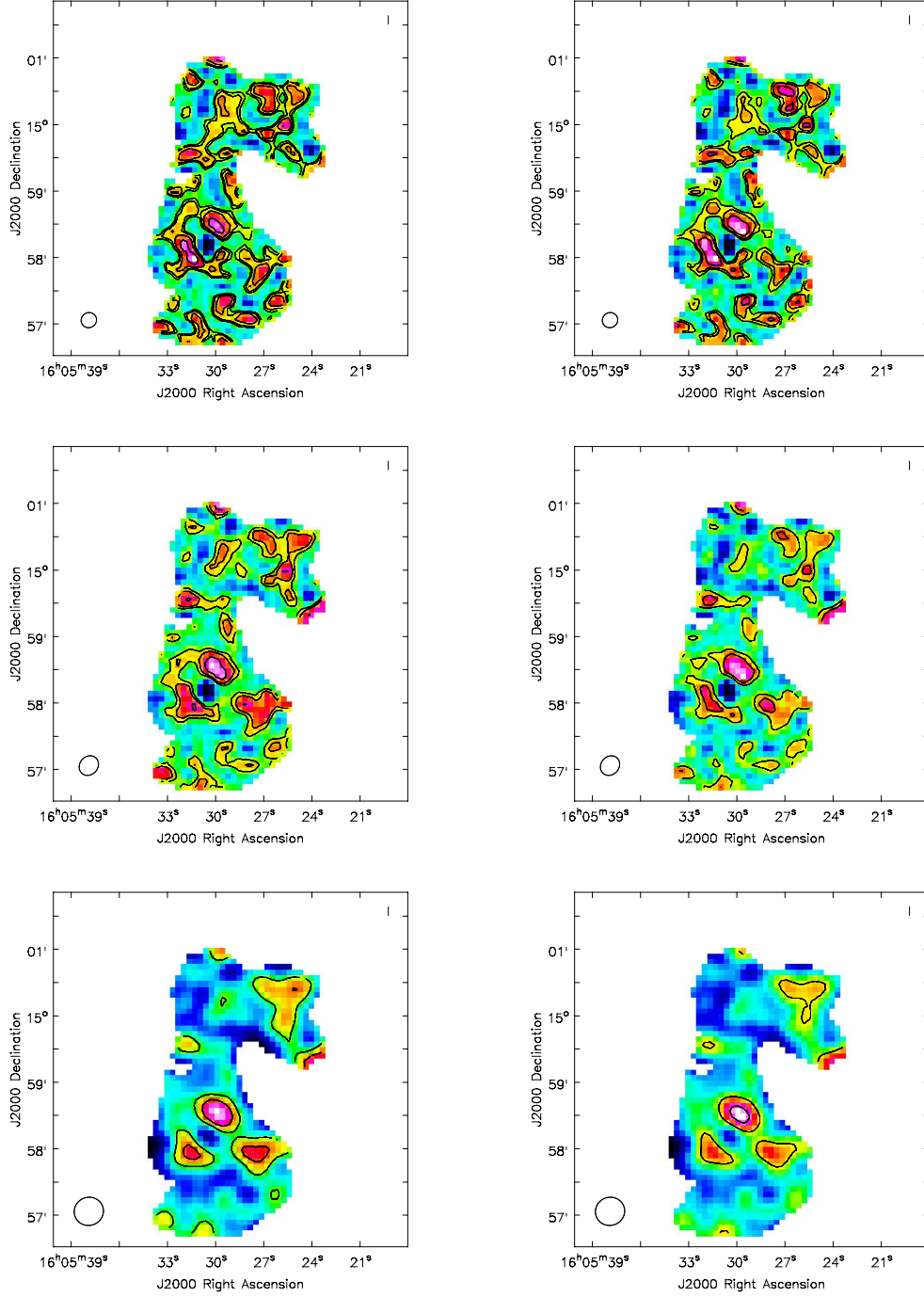


Figure 4.7: Moment zero maps for the robust cube (top row), naturally-weighted cube (middle row), and the tapered cube (bottom row). The left column is the contour maps produced over the widest channel range, and the right column the narrower channel range. In all cases, contours are $[5, 10, 20, 40] \times 10^{19} \text{ atoms cm}^{-2}$ but color scales differ. The effective beam is shown in the lower left of each panel.

higher noise. In fact the flux reported for AGC268069 in the $\alpha.40$ catalog is $0.82 \text{ Jy km s}^{-1}$, differing from the independent flux measurement for the UCHVC catalog of $1.15 \text{ Jy km s}^{-1}$ by almost 30%. This difference in total flux can be seen in Figure 4.8 where the ALFALFA spectrum from $\alpha.40$ is clearly depressed relative to that measured for the UCHVC catalog.

A lower limit to the flux of AGC268069 can be obtained from a single pointed observation with L-band Wide (LBW) at Arecibo, a single pixel receiver. Figure 4.8 also shows the spectrum from the LBW observation. The flux recovered in the LBW observation is $0.62\text{--}0.75 \text{ Jy km s}^{-1}$, depending upon whether the small red wing of emission is included. The difference in including this red wing of emission may account for a substantial fraction of the difference in flux between the two ALFALFA measurements. As a representative total flux of AGC268069, we will use the value of 1 Jy km s^{-1} for comparative purposes to the fluxes detected in the VLA image.

Comparing the LBW observation to the ALFALFA measurement can also provide some gross constraints on the structural parameters of AGC268069. The LBW observation recovers $\sim 60 - 70\%$ of the ALFALFA flux, indicating that the source is relatively compact with the majority of its emission within the central $\sim 4'$. The LBW spectrum shows similar flux levels to the ALFALFA spectrum at the red end of the line, depressed emission in the center of the line, and no emission for the blue end of the line. This may indicate the presence of a large scale velocity gradient from the central pointing of the LBW observation to the outskirts of the source.

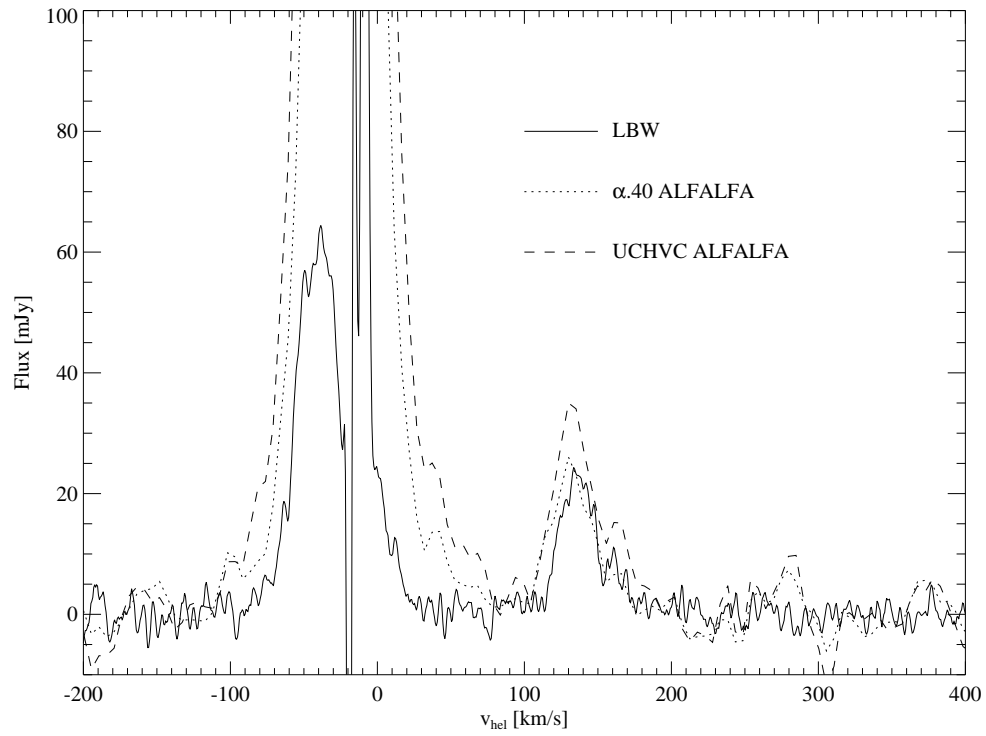


Figure 4.8: The single-dish spectra of AGC268069 for two independent ALFALFA measurements and a single pointed LBW observation.

Recovered VLA Flux

The observed integrated flux densities for the VLA images are calculated by summing all the flux in the final moment zero maps (after correcting for the pedestal level), and the fluxes for both velocity ranges are reported in Table 4.3. As discussed in the section above, the representative single dish integrated flux density for AGC268069 is 1 Jy km s^{-1} . The amount of flux recovered is $\sim 50\%$ for all data cubes. The robust cubes report more flux recovered but are clearly noisier and the extra flux is likely noise that is correlated over the channels used for flux calculation (see Figure 4.6). The flux over the more restrictive channel range is only reduced by $\sim 20\%$ relative to the full channel range. Figure 4.9 shows the robust and natural VLA spectra compared to the ALFALFA and LBW spectra. The VLA spectra are a good match to the LBW spectrum, missing a little bit of the wings of emission. In addition, the blue emission seen in the ALFALFA spectrum that is missed in the LBW observation is also not visible in the VLA observations. This indicates that this emission is not only offset from the center of AGC268069 but also spatially extended or of extremely low surface brightness.

4.5.4 The Velocity Field

The tapered moment zero map for the limited velocity range is clipped at $4 \times 10^{19} \text{ atoms cm}^{-2}$ and used a mask for producing the moment one map, or velocity field, which is also constructed over the limited velocity range. This clipping highlights the three cores in the main component of AGC268069 and the northern extension. Figure 4.10 shows the velocity field of AGC268069 based on the tapered data cube.

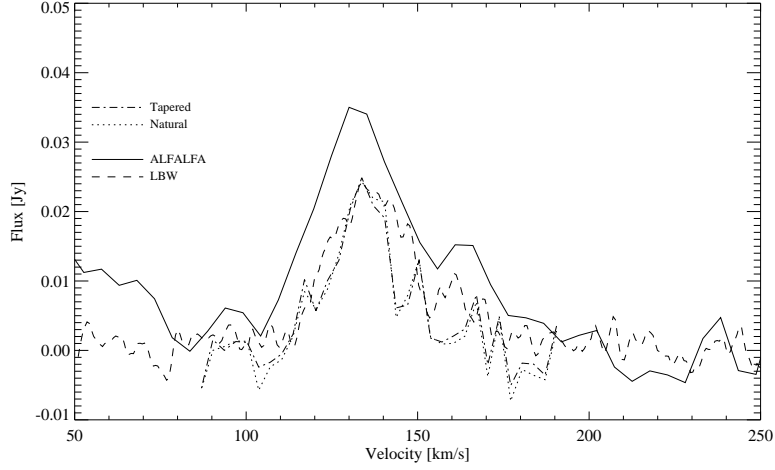


Figure 4.9: The spectrum of AGC268069 from the natural and tapered VLA data cubes. The ALFALFA and LBW spectra are shown for reference.

Two of the cores appear to show tentative evidence for velocity gradients on the order of $4\text{--}5 \text{ km s}^{-1}$, and the northern extension also shows evidence for a velocity gradient.

4.6 Discussion

Based on the exploration above, the tapered data cube at 16kHz resolution is used for exploring the HI structure of AGC268069 and providing constraints on its environment, unless otherwise specifically mentioned.

4.6.1 HI Morphology: Three Cores and the Northern Extension

All the moment zero maps show evidence of three individual cores plus extended northern emission, although these features are seen most clearly in the moment

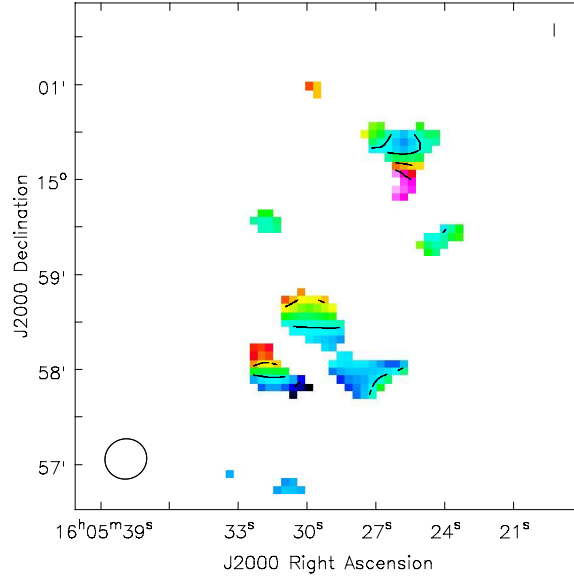


Figure 4.10: The velocity field for the tapered data cube over the velocity range $[120.4, 143.7]$ km s^{-1} , using the moment zero map clipped at $4 \times 10^{19} \text{ atoms cm}^{-2}$ as a mask. Velocity contours are separated by 2 km s^{-1} from 128 to 136 km s^{-1} .

zero maps produced from the tapered data cubes. We refer to these cores by their position throughout this section: the north core, the southeast core, and the southwest core.

The cores are isolated by fitting Gaussian functions using the `imfit` task in CASA. Since the cores are barely resolved and not necessarily well described by a Gaussian, the fits are used to set the central location of each core only. Then, a circular aperture with a radius of $26.5''$ (twice the beam size) is used to extract core properties. The positions and apertures used for the cores are recorded in Table 4.4. The furthest two cores are separated by $\sim 60''$. The northern extended emission is an irregular shape; it is defined interactively using a polygon region in CASA. The central coordinates come from the center of the region, and the extent from the longest axis of the region. It is located $\sim 115''$ ($\sim 2'$) from the northern

Table 4.4. Properties of the HI Cores and Envelope

Name	RA Dec J2000	Aperture ^a "	F_{int}^f ^b Jy km s ⁻¹	F_{int}^i ^c Jy km s ⁻¹	Peak N_{HI} atoms cm ⁻²
North Core	16:05:29.8 +14:58:32	26.5	0.10	0.082	1.7×10^{20}
Southeast Core	16:05:31.5 +14:57:56	26.5	0.061	0.048	1.2×10^{20}
Southwest Core	16:05:27.6 +14:57:53	26.5	0.094	0.062	1.2×10^{20}
Northern Extension	16:05:25.8 +15:00:11	37.5	0.12	0.083	1.0×10^{20}
Envelope	16:05:28.9 +14:59:14	> 30	0.66	—	—

^aRadius of the aperture used to measure source properties. For the Northern Extension, this is an estimate of the largest extent of the irregular mask. For the extended envelope it is a lower limit to the size scale that would be visible in the VLA/C data.

^bIntegrated flux density across the full velocity range [113.7, 153.8] km s⁻¹.

^cIntegrated flux density across the limited velocity range [120.4, 143.7] km s⁻¹.

core. The size and separation of these features is such that they are not resolved by Arecibo 3.5' beam. This can be clearly seen in Figure 4.11, which shows the HI distribution as measured by the VLA compared to the bulk distribution measurable in the ALFALFA HI data.

Table 4.4 also reports the peak column density for the cores and the extension in the tapered data cube. The northern core is the brightest and the two southern cores have identical peak column densities. The northern extension has the lowest peak column density. It is worth noting that the peak column densities of AGC268069 are similar to those seen in C configuration data of Leo P [Giovanelli et al., 2013].

Figure 4.12 shows the spectra of the four features, extracted with the apertures listed in Table 4.4. The north and southeast cores are well detected. The southwest core and northern extension both appear to have elevated baselines. It may be that the pedestaling has spatial variation, or that the elevated flux level seen originally

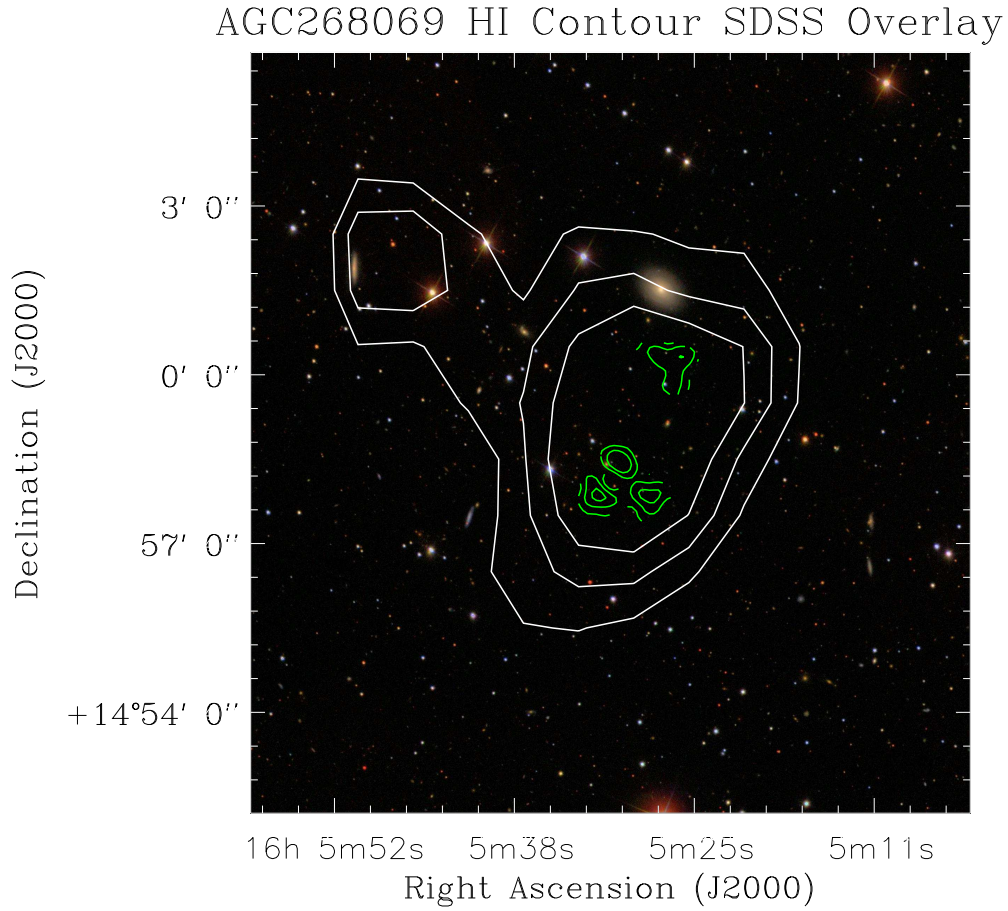


Figure 4.11: Approximate ALFALFA HI contours are shown in white at levels of $[5, 7.5, 10] \times 10^{18} \text{ atoms cm}^{-2}$. The VLA contours for the three cores and northern extension are shown in green; contours are $[5, 10, 20] \times 10^{19} \text{ atoms cm}^{-2}$. The field of view is $\sim 13.5'$, or about 4 ALFALFA beams ($3.5'$ beam). The clumpy emission uncovered by the VLA data is contained within a single ALFALFA beam.

was instead the result of poorly cleaned continuum sources. In either case, the total flux for these two structures is likely an overestimate of the true flux. The north core appears to have a velocity width matching the width of the full source while the southeast core is narrower and concentrated toward the blue side of the total spectra extent. The southwest core is relatively flat spectrally and needs to be examined much more closely for pedestaling issues. The northern extension also appears to be about half the velocity width of the full source, and in this case is concentrated toward the red side of emission.

The fluxes for all four structures are measured from moment zero maps as in the previous section except that the apertures defined in Table 4.4 are used. The four structures are relatively comparable in total flux, although as mentioned above, the fluxes for the southwest core and northern extension are probably elevated relative to their true flux. Together, these four structures recover $\sim 70\%$ of the total flux measured in the tapered data cubes.

There is very little total flux in the emission connecting the four structures, and most of it is on the level of noise. Indeed, about 50% of the total flux for AGC268069 is not recovered in these observations, indicating that there is an extended envelope (on scales $>30''$) containing this missing emission. By assuming that any flux that is not in the four structures discussed is contained in the envelope, limits on the envelope can be placed, and these are discussed in Table 4.4

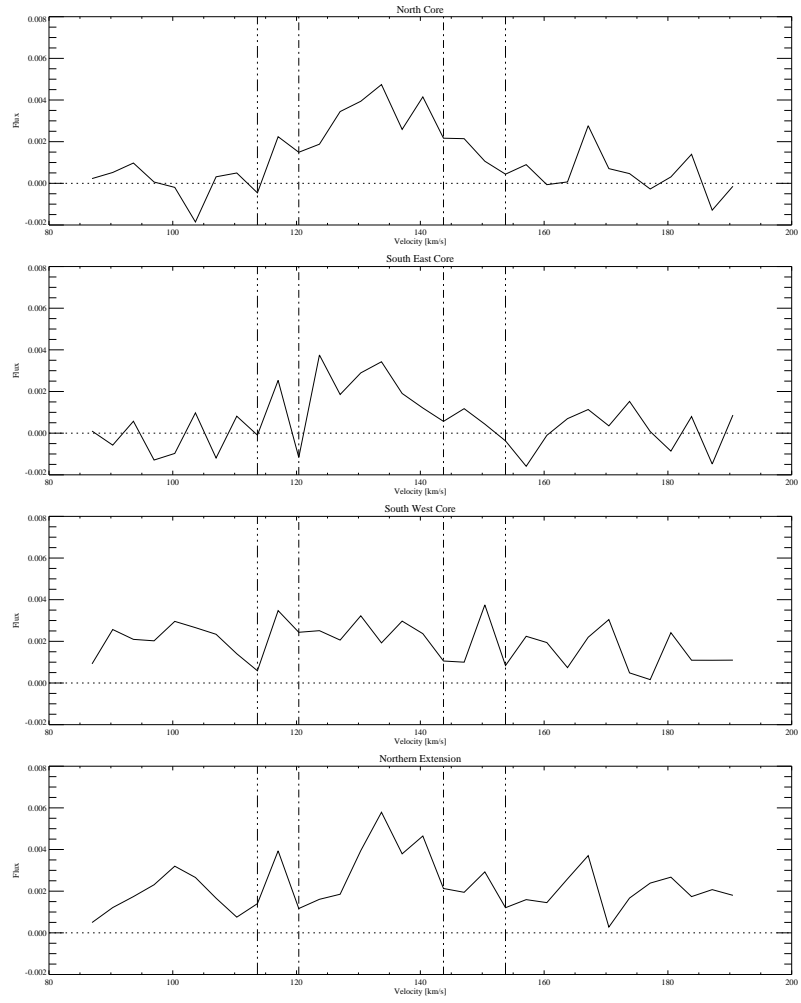


Figure 4.12: The spectra for all three individual cores, plus the northern extension. The vertical lines indicate the velocity range used for creating the moments maps.

4.6.2 The Nature of AGC268069

A clumpy HI structure as seen here is present in both CHVCs [e.g., Brüns and Westmeier, 2004] and in galaxies [e.g., the FIGGS sample of Begum et al., 2006]. Indeed, Leo P shows evidence for a clumpy HI distribution [Bernstein-Cooper et al., 2013]. The presence of cores alone cannot be used to constrain the distance or nature of AGC268069.

Observations of the envelope could help constrain the nature of AGC268069 by being sensitive to the overall morphology and environment. Is the envelope smooth and undisturbed? Or does it show evidence for disruption and interaction with the MW? Is the northern extension of emission potential evidence for interaction? Measuring the spatial extent and kinematics of the envelope is necessary to constrain the total dynamical mass of the system. Understanding the envelope of HI is also necessary for understanding the environment the cores are located in. While we cannot currently constrain the nature of AGC268069, we offer some brief commentary for the circumgalactic and extragalactic possibilities.

AGC268069 as a Circumgalactic Object

At a distance of ~ 250 kpc, AGC268069 would be a circumgalactic object with a total HI mass of $1.5 \times 10^4 M_{\odot}$. The individual core masses would be $\sim 10^3 M_{\odot}$ and the mass of the surrounding envelope $\sim 10^4 M_{\odot}$. Assuming the cores have angular diameters on the order of the beam size, they would be ~ 30 pc in extent.

AGC268069 as an Extragalactic Object

At a distance of ~ 1 Mpc, AGC268069 would be a true extragalactic object separate from the Milky Way. Its HI mass would be $2.4 \times 10^5 M_\odot$ with $1.6 \times 10^5 M_\odot$ in the extended envelope. The cores would have individual masses of $1.4 - 2.4 \times 10^4 M_\odot$ and sizes of 130 pc.

4.7 Improving Constraints

Understanding AGC268069 can be improved by attempting different cleaning and continuum subtraction methodologies to help address the elevated flux level issues. Specifically, I plan to tinker with the parameters of `uvcontsub` and to try continuum subtraction in the image plane using `imcontsub`. I will also produce data cubes with a larger spectral axis and try cleaning to different levels and using a larger clean box to see if that affects the elevated flux level. I will produce spectra over the entire cleaned region and the central box containing the source to constrain any pedestaled flux level before creating a mask for the source, as the elevated flux level affects the clipping used to isolate emission in creating the mask.

Further observations can also help constrain the nature of AGC268069. We have proposed to observe this source (and nine others) at the VLA in the D-configuration. These observations will be sensitive to the extended envelope, allowing us to place more stringent constraints on the environment of AGC268069 and its nature through its HI morphology.

4.8 Summary

High resolution HI observations can be used to indirectly address the hypothesis that UCHVCs are low mass galaxies by using clues in both the HI morphology and kinematics. In addition, high resolution observations are critical for comparison to models, both for the radial distribution of HI and for isolating a CNM component, if any exists.

As a demonstration of how synthesis data can be applied to the study of UCHVCs, one case study is presented here. High resolution observations of AGC268069 reveal the presence of three separate unresolved cores plus extended northern emission. These four individual components recover $\sim 40\%$ of the total flux of AGC268069. Only about 50% of the total flux of AGC268069 is recovered in the VLA data cubes; the missing emission resides in a spatially extended envelope that is resolved out by the array. Observations in a lower resolution configuration (VLA/D) will help recover the missing flux and give clues to its origins by constraining overall HI morphology.

CHAPTER 5

OPTICAL OBSERVATIONS OF ULTRA-COMPACT HIGH VELOCITY CLOUDS

Because Hubble’s Law cannot be used to predict distance from redshift locally, the HI detection of a UCHVC carries no direct distance information. While the minihalo hypothesis for the UCHVCs is intriguing, without a distance constraint it is only a hypothesis. This chapter presents a first report and preliminary discussion of optical follow-up observations of the UCHVCs to search for stellar counterparts and directly address the minihalo hypothesis. The detection of a stellar counterpart immediately confirms that a UCHVC is a galaxy and allows a distance to be constrained via the stellar properties. I first discuss the selection of UCHVCs for optical observations. The observing strategy is then discussed and justified; observations were taken with pODI and DECam, two newly commissioned instruments in shared-risk observing mode. I end by presenting a preliminary reduction and analysis of one target, AGC198606, a compact, high surface brightness source located near Leo T. This work is part of an ongoing project to analyze optical data of twenty-nine systems.

5.1 Target Selection

Follow-up observations require the dedication of large amounts of telescope time, and we wish to select the best sample of galaxy candidates for the optical follow-up work. In order to do this, an expanded search space (spatially and spectrally) was considered for the UCHVCs and a set of criteria were defined for selecting the

best sources based on isolation, compactness, column density, recessional velocity, and association with known galaxies.

5.1.1 The Sky Searched

As ALFALFA data continue to be processed, more and more of the “ALFALFA sky” becomes available. At the time of the optical observations (spring 2013), the available low velocity ALFALFA spring sky (R.A.: $7.5h - 16.5h$; $cz < 3000 \text{ km s}^{-1}$) was $0^\circ - 32^\circ$ in declination, an increase of $\sim 100\%$ in spring coverage from the $\alpha.40$ footprint considered in Chapter 3. In addition, the velocity constraints are relaxed so that the full velocity space between $[-500, 1000] \text{ km s}^{-1}$, including velocities of Galactic HI, is considered, although the isolation of these sources is difficult to determine and most of them were identified through visual inspection rather than in any automated way.

5.1.2 Selecting the Best Galaxy Candidates

All the targets for optical observations meet the criteria of UCHVCs presented in Chapter 3 (except for the allowance of Galactic velocities); however, we wish to select subsamples that are the most likely to be galaxy candidates. Several criteria are used to select the best candidates.

Both Leo T and Leo P have high column densities compared to the general population of UCHVCs (and HVCs). A high column density of HI is also expected as a requisite for star formation. Therefore, we wish to target the clouds

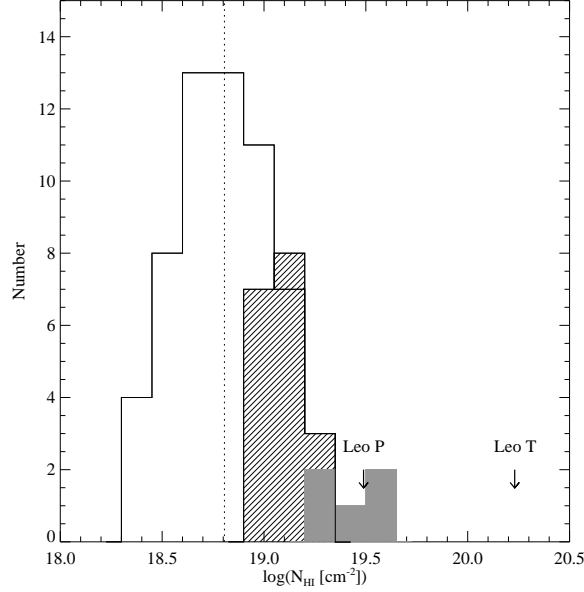


Figure 5.1: The distribution of \bar{N}_{HI} for the $\alpha.40$ UCHVCs (solid line), the UCHVCs with the highest \bar{N}_{HI} values ($\bar{N}_{HI} \geq 10^{19.2}$; filled histogram) and the UCHVCs with high \bar{N}_{HI} values ($10^{19} \leq \bar{N}_{HI} < 10^{19.2}$; hashed histogram). The median value of the $\alpha.40$ UCHVCs is shown by the dotted vertical line. The values of Leo P and Leo T from the ALFALFA data are shown for comparison.

with the highest \bar{N}_{HI} values as good minihalo candidates. Figure 5.1 compares a selection of sources with extreme \bar{N}_{HI} values to the full distribution of the $\alpha.40$ UCHVC catalog. Note that not all the UCHVCs considered here are part of the $\alpha.40$ catalog as a result of the expanded coverage (spatial and kinematic) considered. The highest \bar{N}_{HI} sources are shown in the solid histogram; they have $\bar{N}_{HI} \geq 10^{19.2}$ atoms cm^{-2} . Sources with $\bar{N}_{HI} \geq 10^{19}$ atoms cm^{-2} (and $< 10^{19.2}$) are shown by the hashed histogram. The highest \bar{N}_{HI} sources have values similar to that measured for Leo P in the ALFALFA dataset and the high \bar{N}_{HI} sources lie above the median value of \bar{N}_{HI} for the $\alpha.40$ UCHVC catalog.

Leo T and Leo P are also both very compact objects and are unresolved in the ALFALFA data. In addition, models for gas in low mass dark matter halos predict

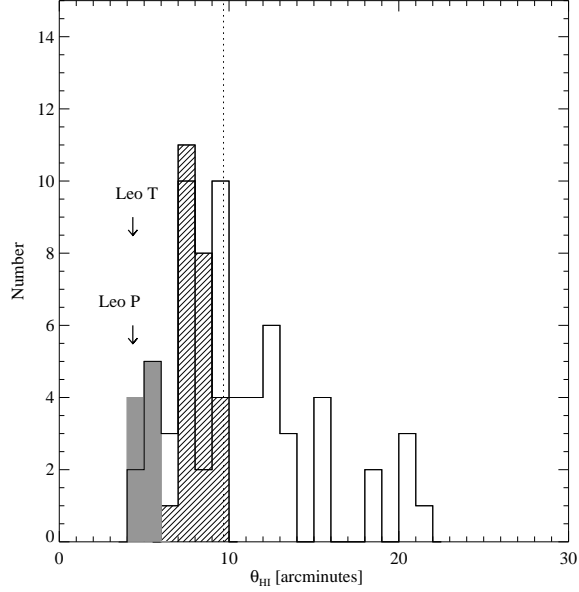


Figure 5.2: The distribution of angular diameters for the $\alpha.40$ UCHVCs (solid line), the “most compact” UCHVCs ($\theta_{HI} \leq 6'$; filled histogram) and the “compact” UCHVCs ($6' < \theta_{HI} \leq 10'$; hashed histogram). The median value of the $\alpha.40$ UCHVCs is shown by the dotted vertical line. The values of Leo P and Leo T from the ALFALFA data are shown for comparison.

HI diameters on the order of $0.2 - 0.5$ kpc, or $\sim 1 - 3'$ for a range of reasonable distances for the objects [Faerman et al., 2013]. Thus, the most compact clouds are among the best galaxy candidates. Although the UCHVCs are selected, by definition, to be compact, we wish to focus on the most compact subset as the best candidates for detection of a stellar counterpart in optical observations. Figure 5.2 shows the distribution of HI angular diameters (θ_{HI}) for the $\alpha.40$ UCHVCs. UCHVCs with $\theta_{HI} \leq 10'$ are below the median size of the $\alpha.40$ UCHVCs; we refer to these objects as “compact” UCHVCs. In addition, UCHVCs with $\theta_{HI} \leq 6'$ are defined as the “most compact” UCHVCs and are among the best galaxy candidates.

Given the lack of distance information to the UCHVCs, it must always be con-

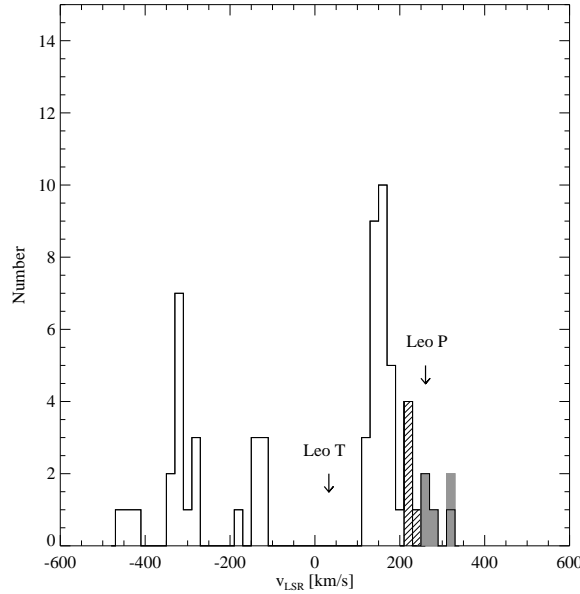


Figure 5.3: The distribution of v_{LSR} for the $\alpha.40$ UCHVCs (solid line), the UCHVCs with the highest v_{LSR} values ($v_{LSR} \geq 250 \text{ km s}^{-1}$; filled histogram) and the UCHVCs with high v_{LSR} values ($200 \text{ km s}^{-1} \leq v_{LSR} < 250 \text{ km s}^{-1}$; hashed histogram). The values of Leo P and Leo T from the ALFALFA data are shown for comparison.

sidered that the UCHVCs arise from Galactic origins. As discussed in Chapter 3, it is hard to account for large recessional velocities in Galactic halo models of UCHVCs. Hence, clouds with large recessional velocities are good candidates for optical follow-up. Figure 5.3 compares the distribution of LSR velocities for the $\alpha.40$ UCHVCs to the sample of clouds with extreme positive velocities targeted for optical follow-up. We note that the distribution of $\alpha.40$ UCHVCs includes clouds with large negative recessional velocities; these are velocities of clouds associated with the Magellanic Stream (fall sky) and do not represent good extragalactic candidates. Clouds with $v_{LSR} \geq 250 \text{ km s}^{-1}$ (solid histogram) and those with $200 \text{ km s}^{-1} \leq v_{LSR} < 250 \text{ km s}^{-1}$ (hashed histogram) are unlikely to be explained by Galactic processes. For reference, Figure 5.4 shows how the UCHVCs with ex-

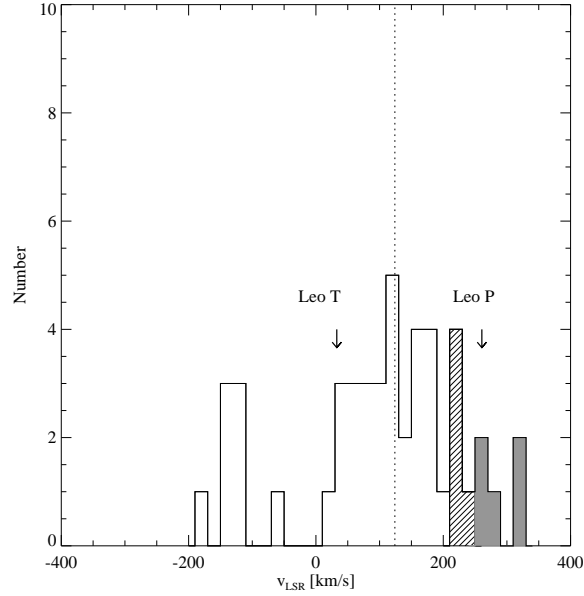


Figure 5.4: The distribution of v_{LSR} for all UCHVCs considered as optical targets (solid line), the UCHVCs with the highest v_{LSR} values ($v_{LSR} \geq 250 \text{ km s}^{-1}$; filled histogram) and the UCHVCs with high v_{LSR} values ($200 \text{ km s}^{-1} \leq v_{LSR} < 250 \text{ km s}^{-1}$; hashed histogram). Leo T and Leo P are shown for reference. The median value of the full UCHVC optical sample is shown by the dotted line.

treme velocities compare to the full sample of UCHVCs considered for optical follow-up. Here it is clear that the UCHVCs with extreme velocities represent a tail-end of the distribution and lie well above the median recessional velocity of the full sample.

UCHVCs that are the most likely to represent gas in dark matter halos are expected to be isolated from other HVC emission; UCHVCs near other HVC structure are likely associated with that structure and arise from a Galactic or circumgalactic phenomenon. For this reason, UCHVCs that meet the definition of being in the most-isolated subsample (MIS) defined in Chapter 3 are considered to be the best candidates for optical follow-up.

We also give added weight to UCHVCs that have a location near known galaxies (or star clusters), suggesting that they might arise from a similar origin. For example, HVC214.78+42.45+47 (AGC198606) is separated from Leo T by $70'$ and 13 km s^{-1} and has twice its HI flux density. HVC240.22+41.75+10 (AGC501816) is separated by $7'$ and 40 km s^{-1} from Sextans C, a distant globular cluster. HVC274.68+74.70-123 (AGC226067) is located near several galaxies that are associated with subcluster A of the Virgo cluster [Binggeli et al., 1993]; it may be a dark galaxy at the distance of Virgo. HVC314.57+74.80+218 (AGC238626) is separated from the dwarf galaxy GR8 by 2.4° (a linear separation of 92 kpc at the distance of GR8) and is at essentially the same velocity.

5.1.3 The Targets

In Table 5.1, we present the targets selected for optical follow-up on the basis of the criteria discussed above. The columns of the table are as follows:

- Column 1: Source name in the traditional format for HVCs.
- Column 2: Identification number in the Arecibo General Catalog (AGC), an internal database maintained by MH and RG, included to ease cross-reference with our archival system and the α .40 catalog.
- Column 3: Velocity in the local standard of rest frame for an assumed solar motion of 20 km s^{-1} towards $l = 57^\circ, b = 25^\circ$.
- Column 4: HI line full width at half maximum (W_{50}).
- Column 5: Average HI angular diameter.
- Column 6: HI mass for an assumed distance of 1 Mpc.

- Column 7: Average column density of the UCHVC from Equation 3.6 in Chapter 3.
- Column 8: S/N of the detection.
- Column 9: Number of HVC neighbors within $D = 3^\circ$ from the definition in Chapter 3 and for the expanded sky coverage considered in this chapter.
- Column 10: Number of HVC neighbors within $D = 10^\circ$ with the same qualifications as above.
- Column 11: This column contains notes on the sources selected for optical imaging. ‘MIS’ refers to the sources that meet the MIS criteria defined in Chapter 3 and are well separated from the Galactic HI with $|v_{LSR}| > 120$ km s⁻¹. The UCHVCs with $\bar{N}_{HI} \geq 10^{19.2}$ atoms cm⁻² are indicated by “highest \bar{N}_{HI} ” and those with $10^{19} \leq \bar{N}_{HI} < 10^{19.2}$ by “high \bar{N}_{HI} ”. The sources with $\theta_{HI} \leq 6'$ are labeled “most compact” and those with $6' < \theta_{HI} \leq 10'$ “compact”. Sources with $v_{LSR} \geq 250$ km s⁻¹ are indicated with “highest v_{LSR} ” and those with $200 \text{ km s}^{-1} \leq v_{LSR} < 250 \text{ km s}^{-1}$ with “high v_{LSR} ”.

There are four sources included in Table 5.1 which were selected as good candidates based on HI properties but were removed from consideration for optical follow-up based on the presence of bright foreground stars that would contaminate any images. These four sources are marked with a note.

Table 5.1. ALFALFA UCHVCs Targeted for Optical Observations

Source	AGC	v_{LSR} km s ⁻¹	W ₅₀ km s ⁻¹	θ '	M _{H I} ^a M _⊙	\bar{N}_{HI} cm ⁻²	S/N	N ₃	N ₁₀	Notes
HVC205.28+18.70+150	174540	150	23	8	5.7	19.2	28.0	0	3	high \bar{N}_{HI} , compact, MIS
HVC193.11+28.88+174 ^b	189177	174	22	9	5.3	18.7	9.3	1	23	compact
HVC214.78+42.45+ 47	198606	47	21	9	6.2	19.6	87.0	1	5	highest \bar{N}_{HI} , compact
HVC204.88+44.86+147	198511	147	15	7	5.2	18.8	14.0	0	3	compact, MIS
HVC205.83+45.14+173	198683	173	19	10	5.3	18.6	9.5	2	3	MIS
HVC215.28+49.73+ 79	198693	79	22	16	6.0	18.8	23.9	0	3	
HVC240.22+41.75+101	501816	101	16	7	5.6	19.2	17.2	0	28	high \bar{N}_{HI} , compact
HVC215.57+55.63- 62	208752	-62	16	8	5.6	19.1	25.0	0	2	high \bar{N}_{HI} , compact
HVC217.77+58.67+ 96	208747	96	23	11	5.8	19.0	27.6	1	15	high \bar{N}_{HI}
HVC248.02+50.88+174	208524	174	14	5	5.2	19.1	11.6	0	12	high \bar{N}_{HI} , most compact
HVC212.68+62.39+ 64	208753	64	23	13	6.0	19.0	28.9	0	2	high \bar{N}_{HI}
HVC221.35+65.75+127 ^b	219655	127	25	14	5.7	18.7	14.3	1	13	
HVC197.61+70.83+ 93	219662	93	21	11	5.4	18.6	8.6	0	6	
HVC230.27+71.10+ 76	219663	76	17	7	5.3	18.8	15.1	0	5	compact
HVC245.26+69.53+217	215417	217	17	10	5.2	18.5	8.8	0	3	high v_{LSR} , compact, MIS
HVC235.38+74.79+195	219656	195	21	8	5.3	18.8	12.5	1	4	compact, MIS
HVC277.25+65.14-140	227977	-140	23	5	5.0	18.9	8.0	0	2	most compact, MIS
HVC274.68+74.70-123	226067	-123	54	5	5.3	19.3	11.0	0	0	highest \bar{N}_{HI} , most compact, MIS
HVC271.57+79.03+248	229326	248	23	8	5.3	18.7	9.6	1	6	high v_{LSR} , compact
HVC276.53+79.84+255	229327	255	19	11	5.3	18.5	9.4	1	4	highest v_{LSR} , MIS
HVC298.95+68.17+270	227987	270	26	12	6.1	19.2	44.2	0	4	highest v_{LSR} , high \bar{N}_{HI} , MIS
HVC299.62+67.65+326	227988	326	41	5	5.1	18.9	6.4	0	0	highest v_{LSR} , most compact, MIS
HVC314.57+74.80+218	238626	218	36	4	4.9	19.0	5.2	0	17	high v_{LSR} , high \bar{N}_{HI} , most compact
HVC333.71+75.23+118	239116	118	24	7	5.2	18.7	9.7	2	11	compact
HVC330.13+73.07+132	233831	132	16	4	5.2	19.2	10.8	0	12	high \bar{N}_{HI} , most compact

5.2 The Observations

In order to search for a stellar counterpart, the UCHVCs in the optical sample were observed in the *g* and *i* filters to depths of ~ 25 th magnitude. This approach enables a two filter search for an underlying old population based on detection of the red giant branch (RGB) out to distances of ~ 1.5 Mpc. The *i* filter was chosen for sensitivity to the RGB population. In addition, a third filter, *r*, was obtained for a few sources to help with star-galaxy discrimination based on color. These sources will be used to test if the increase in detection efficiency offsets the extra observing time required. As all the spring UCHVC targets lie in SDSS coverage, no standards were observed and photometry is bootstrapped from SDSS photometry (see Section 5.4.1).

In the spring of 2013, three observing programs were dedicated to these observations. The first of these programs was four nights on the WIYN 3.5m telescope with pODI (March 14-17) awarded to E. A. K. Adams through NOAO time on the telescope. The second program was also four nights on the WIYN 3.5m telescope with pODI (Apr 12-15) awarded to J. J. Salzer through the allocation at Indiana University. The last program was 3 nights with DECam at CTIO (Apr 28-30) awarded to R. Muñoz through the time granted to Chilean institutions. Since both instruments were available through shared risk observing and the software required to perform initial pipeline processing is not fully developed, here we present only a few demonstrative results. Tables 5.2, 5.3 and 5.4 summarize the sources observed as part of each program.

Table 5.1 (cont'd)

Source	AGC	v_{LSR} km s ⁻¹	W ₅₀ km s ⁻¹	θ '	M_{HI}^a M_{\odot}	\bar{N}_{HI} cm ⁻²	S/N	N_3	N_{10}	Notes
HVC336.23+75.97+114 ^b	232765	114	23	5	5.5	19.4	25.9	1	9	highest \bar{N}_{HI} , most compact
HVC326.91+65.25+316	238713	316	26	11	5.5	18.7	10.9	0	0	highest v_{LSR} , MIS
HVC346.46+62.52+ 69	249320	69	24	5	5.3	19.1	11.2	0	2	high \bar{N}_{HI} , most compact
HVC340.05+57.54-122	249441	-122	16	8	5.2	18.7	9.7	0	4	compact, MIS
HVC 3.96+69.03+ 78	249000	78	31	9	5.6	19.0	20.8	0	4	high \bar{N}_{HI} , compact
HVC 28.09+71.87-142	249393	-142	38	13	5.4	18.4	7.6	0	0	MIS
HVC 11.76+67.89+ 60	249525	60	24	9	6.2	19.6	77.5	0	3	highest \bar{N}_{HI} , compact
HVC353.41+61.07+257	249323	257	20	11	5.5	18.7	13.0	3	4	highest v_{LSR} , MIS
HVC353.89+61.13-135	249324	-135	21	8	5.4	18.9	15.0	0	16	compact
HVC351.17+58.56+214	249282	214	40	6	5.5	19.3	17.4	0	4	high v_{LSR} , highest \bar{N}_{HI} , most compact, MIS
HVC 15.96+63.90+ 44	249565	44	18	8	5.6	19.1	23.5	0	1	high \bar{N}_{HI} , compact
HVC 25.54+59.19+ 42 ^b	258682	42	25	9	5.6	19.0	14.5	0	0	high \bar{N}_{HI} , compact
HVC 5.58+52.07+163	258459	163	24	10	5.5	18.7	12.9	0	4	MIS
HVC 13.59+54.52+169	258237	169	23	7	5.5	19.1	17.1	1	4	high \bar{N}_{HI} , compact, MIS
HVC 13.60+54.23+179	258241	179	17	10	5.4	18.7	10.6	1	5	compact
HVC 13.63+53.78+222	258242	222	21	7	5.2	18.8	8.7	0	1	high v_{LSR} , compact, MIS
HVC 24.33+53.49+ 21	258683	21	14	9	5.6	19.0	20.9	0	0	high \bar{N}_{HI} , compact
HVC 28.07+43.42+150	268069	150	29	7	5.4	19.0	10.8	0	11	high \bar{N}_{HI} , compact
HVC 28.03+41.54+127	268071	127	62	9	5.8	19.1	17.7	1	9	high \bar{N}_{HI} , compact
HVC 19.13+35.24-123	268213	-123	17	11	5.8	19.0	22.4	0	18	high \bar{N}_{HI}
HVC 27.86+38.25+124	268074	124	23	10	5.5	18.8	12.6	2	4	compact, MIS

^aFor an assumed distance of 1 Mpc.

^bSources removed from planned optical observations due to contamination by bright foreground stars.

Table 5.2. ALFALFA UCHVCs Observed in March WIYN Run

Source	Filters	t_{tot}^a sec	Notes
AGC174540	g, i	9×300	Seeing 0.7''. First 15 minutes of i in twilight
AGC198606	g, i	9×300	Seeing 0.6 – 0.7''.
AGC198511	g, i	9×300	Seeing 0.8 – 1''.
AGC198683	g, i	9×300	Seeing 0.8 – 1''. Patchy clouds, lost guide star a few times in g.
AGC208747	g, i	9×300	Seeing 0.5''. Clouds for first 10-15 minutes in i.
AGC215417	g, i	9×300	Seeing 0.7 – 1''.
AGC226067	g, i	9×300	Seeing 0.7''.
AGC219656	g, i	9×300	Seeing 0.7''.
AGC227987	i	9×300	Seeing > 1''.
AGC238626	g, i	9×300	Seeing 0.5 – 0.7''.
AGC238713	g, i	9×300	Seeing 0.6 – 1''. Light clouds for last 10-15 minutes in g.
AGC249525	g, i	9×300	Seeing 0.5 – 0.7''.
AGC249282	g, i	9×300	Seeing 0.6 – 0.8''.
AGC258237	i	9×300	Picked up at end of nights
AGC258242	g, i	9×300	Seeing 0.7 – 0.8''.
AGC268069	i	9×300	Seeing 0.5 – 0.7''. Last 10 minutes in twilight

^aTotal time spent on source; effective source time will be less due to the dither pattern to account for cell and chip gaps.

Table 5.3. ALFALFA UCHVCs Observed in the April WIYN Run

Source	Filter	t_{tot}^a sec	Notes
AGC198683	g	9×300	Seeing 1.0-1.2''. Replacement for cloudy exposures from first night.
AGC208753	g, i	9×300	Seeing 0.9 – 1.1''.
AGC229326	g, i	9×300	Seeing 0.9 – 1.2''.
AGC249323	g, i	9×300	Seeing 0.9 – 1.2''.
AGC258237	g	9×300	Seeing 1.0 – 1.1''. i observed during previous run
AGC198606	r	9×300	Seeing 0.9'' and up. Third filter for color discrimination.
AGC229327	g, i	9×300	Seeing 0.5 – 0.9'' for i; 1.1 – 1.4'' and dubious focus for g.
AGC268069	g	9×300	Seeing 1.1 – 1.4''. i observed during previous run
AGC249393	g, i	9×300	Seeing 0.9 – 1.6''.
AGC226067	r	9×300	Seeing 0.9 – 1.4''. Third filter for color discrimination.
AGC249565	g, i	9×300	Seeing 1.4 – 1.8''.
AGC219663	g, i	9×300	Seeing 1.2 – 1.6''.
Leo P	g, i	9×300	Seeing 1.0 – 1.3''. Comparison and test for methodology.

^aTotal time spent on source; effective source time will be less due to the dither pattern to account for cell and chip gaps.

Table 5.4. ALFALFA UCHVCs Observed in the April CTIO Run

Source	Filters	t_{tot} sec	Notes
AGC249324	g, i	3×600	Clouds in the area; i frames probably useless.
AGC258241	g, i	3×600	Slightly better conditions. Seeing 1".4
AGC268071	g, i	3×600	Seeing 1".8
AGC268213	g, i	3×600	Clouds in the area affect i frame significantly. Seeing 1".6
AGC501816	g, i	3×600	Seeing 1".5
AGC227988	g, i, r	3×600	Seeing 1".5. Clouds near the area.
AGC249323	g, i	3×600	Seeing 1".3. Clouds near the area
AGC258242	g, i, r	3×600	Seeing 1".5. Clouds near the area
AGC268074	g, i	3×600	Seeing 2".0. Clouds near the area

5.3 Data Reduction and Calibration

This section contains an overview of the reduction and calibration of the data associated with one UCHVC, AGC198606. The data were obtained with pODI during the March run. A full accounting of the reduction and calibration is given in Appendix C. As detailed earlier in this chapter, AGC198606 is a compact source and is among the highest \bar{N}_{HI} sources. In addition, it is located spatially and kinematically near Leo T.

5.3.1 pODI: The “partially-filled” One Degree Imager

pODI is an early deployment of the One Degree Imager (ODI) for the WIYN 3.5m telescope. It has a fraction of the orthogonal transfer arrays (OTAs) needed to fill the full focal plane of ODI. pODI consists of an array of 3×3 OTAs providing a $24' \times 24'$ field of view, plus an additional four outlying OTAs which sample the full focal plane and are used for guiding. Due to gaps between the OTAs and

gaps within the OTAs from the individual cells, there is a standard 9-point dither pattern for observations with pODI. This dither pattern can be executed from a single observation call; during the March run the dither pattern was broken into three separate observing calls to allow for natural breaks to check focus and to deal with any potential telescope problems.

The basic data product from pODI is the single OTA as these are independent chips which are calibrated separately. Hence each exposure has nine science OTAs (plus four ancillary OTAs), and a full nine-point dither pattern has, at a minimum, 81 OTAs that need to be calibrated and combined.

5.3.2 The Standard Calibration Procedure

pODI has a standard calibration procedure for obtaining calibration data at the telescope in order to facilitate pipeline processing of the data. The standard calibration data include afternoon calibrations consisting of biases, dark frames, and dome flats in the Sloan griz filters, plus any additional filters the observer plans on using. During evening twilight, short observations are taken of a SDSS calibration field in the griz filters to monitor photometric and seeing quality at the site.

5.3.3 Pipeline Processing

The pODI data are processed by a pipeline being developed at NOAO. As of June 2013, the data is run through the pipeline at NOAO and 1 – 2 weeks after an observing run, the observer receives an email with directions on how to access the

calibrated data at an FTP site and a brief description of the data products. The pipeline is being transferred to Indiana University where it will soon be available to the observer to run through the ODI Portal, Pipeline and Archive (PPA)¹ interface. Currently, only raw data is generally available through the PPA.

The current version of the pipeline does the bias and dark current subtraction and the flat field correction. It does not perform a dark sky or fringe correction, account for persistency from saturation, or remove cosmic rays. Ghost pupil corrections are done for calibration flat images but not science images.

The pipeline does apply a world coordinate system to all frames and reprojects all the individual OTAs onto a common frame. The pipeline both stacks all the frames within a given observe command (can consist of multiple exposures from a dither pattern) and identifies exposures with a common pointing (and filter) and stacks those objects. The final products produced by the pipeline include:

- sframes: single OTA frames with basic data reduction applied
- rframes: single OTA frames reprojected onto a common coordinate system
- dstacks: stacked images from a dither pattern executed within a single observing call
- pstacks: stacked images from a common pointing reference

Issues with the Pipeline

The pipeline is currently under development and hence does not perform perfectly. During the March run, twilight flats were obtained in addition to the stan-

¹<http://portal.odi.iu.edu/index/front>

dard dome flats; the intent was to use the twilight flats to improve the data calibration if necessary. The pipeline did not distinguish between the dome and sky flats and incorrectly chose to combine the two different types of flat field images. A single night of data (the first night) was rerun through the pipeline at Indiana University with only dome flats used for flat fielding.

In addition, there is concern about how the pipeline does scaling, especially for the sky background when combining different images (D. Harbeck, private communication 2013). For the preliminary analysis presented below, the combined images produced by the pipeline were used for expediency but it should be noted that a more careful combination and stacking by the observer (perhaps using **swarp**) will be necessary for fully accurate results.

5.4 Post-Pipeline Processing

While the work presented here will need to be redone after the next pipeline release, the following sections illustrate the processing steps and their preliminary results.

The pipeline-stacked images were manually edited to remove cosmic rays and satellite trails, although many cosmic rays still escaped this step. In addition, the images were shifted to be aligned with each other so that for following analysis pixel coordinates were consistent between the two filters.

5.4.1 Photometric Calibration

The pipeline does attempt to produce photometric calibration solutions for the observed fields but there is not yet any documentation for how the photometry is derived and the zero points recorded in the header are observed to be inaccurate. Since all the fields observed for this project are within the SDSS footprint and the UCHVCs were observed in Sloan filters, relative photometry for each field is performed using SDSS stars in that field.

A list of SDSS stars within a $10'$ radius of the field center and $17 < i < 20$ and $17 < g < 20$ was generated through the SDSS SkyServer Search Forum². These stars were then overlaid on the i broadband image in IRAF³. Approximately 20 stars were identified in the image that appeared to be relatively isolated and neither saturated nor too faint. These stars were then overlaid on the g image and those that passed visual inspection in both filters were used for photometry.

The magnitudes of the selected stars were measured with an aperture of 25 pixels (or $2.75''$) in radius, about 4 times the FWHM of the point spread function (PSF) of the stars in the image. Comparing the measured magnitudes from the images to the SDSS magnitudes, the photometric zero-point and color terms were obtained. For i , the zero point is 25.83 with a color term of 0.0027 and a photometric scatter of 0.015; for g these values are 26.23 and 0.11 with a scatter of 0.029. These values are consistent with those found by S. Janowiecki for pODI data (private communication 2013), verifying the photometric solutions. The color term for g is as expected given that the filter is not completely identical to the Sloan

²<http://cas.sdss.org/dr7/en/tools/search/form/form.asp>

³IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under cooperative agreement with the National Science Foundation.

filter. Full details of the photometric calibration are contained within Appendix C.2.

5.5 Analysis of the Stellar Population

The general methodology for identifying stellar counterparts within the optical data follows that of Walsh et al. [2009]. Briefly, the steps we follow are:

- Identify the stars within an image and obtain their magnitudes.
- Filter the stars based on the expected stellar population and a given distance.
- Spatially smooth the filtered stars to search for evidence of an overdensity.
- Repeat the two previous steps for a range of different distances.

5.5.1 Identify and Measure Stars

We identified the stars with an image and obtained their photometry using **daophot** [Stetson, 1987]; the steps undertaken followed the **daophot** cookbook for IRAF [Davis, 1994], and full details are available in Appendix C.3. Briefly, stars were identified with `daofind` in a single filter (i). Due to the presence of cosmic rays, a relatively high detection threshold was set and the identified sources were inspected by eye and many were removed from the source list. In addition, the g image was examined to ensure that the sources were visible in both filters; otherwise they were removed from the photometry list.

After the stars were identified, aperture photometry with a small aperture (6.5 pixels or $\sim 0.7''$, approximately the FWHM) was performed. This served as a starting point for the PSF-fitting based photometry and also allowed a determination of sky levels since no sky subtraction was performed. The small aperture is necessary to prevent blending of crowded objects. A model of the PSF was created using a handful of isolated, well-behaved stars. The photometry of all the stars was then obtained by running `allstar` in IRAF which fit a PSF to each identified star and subtracted it from the image. No new stars were identified at this stage, so the photometry list produced by this step was taken as the final catalog. As the photometry for these stars was based on a small aperture, an aperture correction of 0.3 magnitudes was calculated by comparing the photometry results from the PSF-fitting method to large aperture photometry (25 pixels, the aperture used for determining photometric solutions).

5.5.2 Filtering the Stars

Following Walsh et al. [2009], the stars in the field are filtered based on a two-age, metal poor stellar population. The Girardi et al. [2004] isochrones for ages of 8 and 13.5 Gyr and $[\text{Fe}/\text{H}]$ of -1.5 and -2.27 are used to define a filter for the stars based on a i vs $g - i$ CMD. The filter is widened by adding an uncertainty of 0.2 magnitudes in the color determination. As the minihalo hypothesis for the UCHVCs places them at distances beyond the detection limits of SDSS, the distance range considered for the various filters is 300-1500 kpc; specifically we use a range of distance moduli from 22.4 to 26 separated by 0.4. Figure 5.5 shows the filter overlaid on the CMD of the full field of AGC198606, our test case, for the two

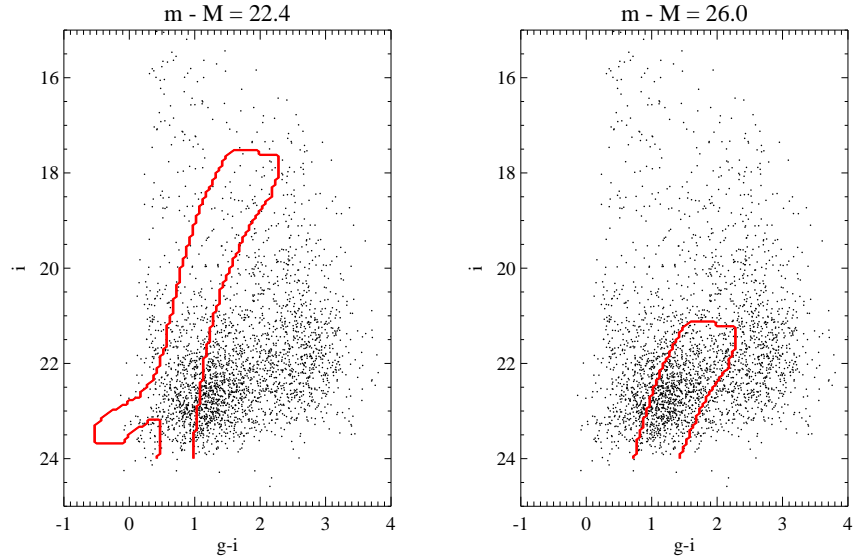


Figure 5.5: The filter for the two extremes of the distance range considered (300 kpc, left; 1500 kpc, right) overlaid on the CMD for all objects identified within the AGC198606 frame.

bounding distance moduli of 22.4 and 26. Note that due to background issues and the high threshold used in detecting sources that the true sensitivity of the images is certainly much deeper than that seen in Figure 5.5.

5.5.3 Smoothing the Filtered Population

After the stars identified within the field are filtered, the stars are spatially smoothed to scales of $1'$, $2'$, $3'$, and $4'$. Walsh et al. [2009] smoothed to scales of $4'$ only based on observed sizes of Local Group dwarf galaxies. Given that the search for a stellar overdensity is more focused here, it is computationally reasonable to smooth to multiple scales. At presumed further distances than the UFDs of SDSS, the stellar counterparts to the UCHVCs may have smaller angular sizes and so smoothing on smaller scales may help reveal a stellar population. For example,

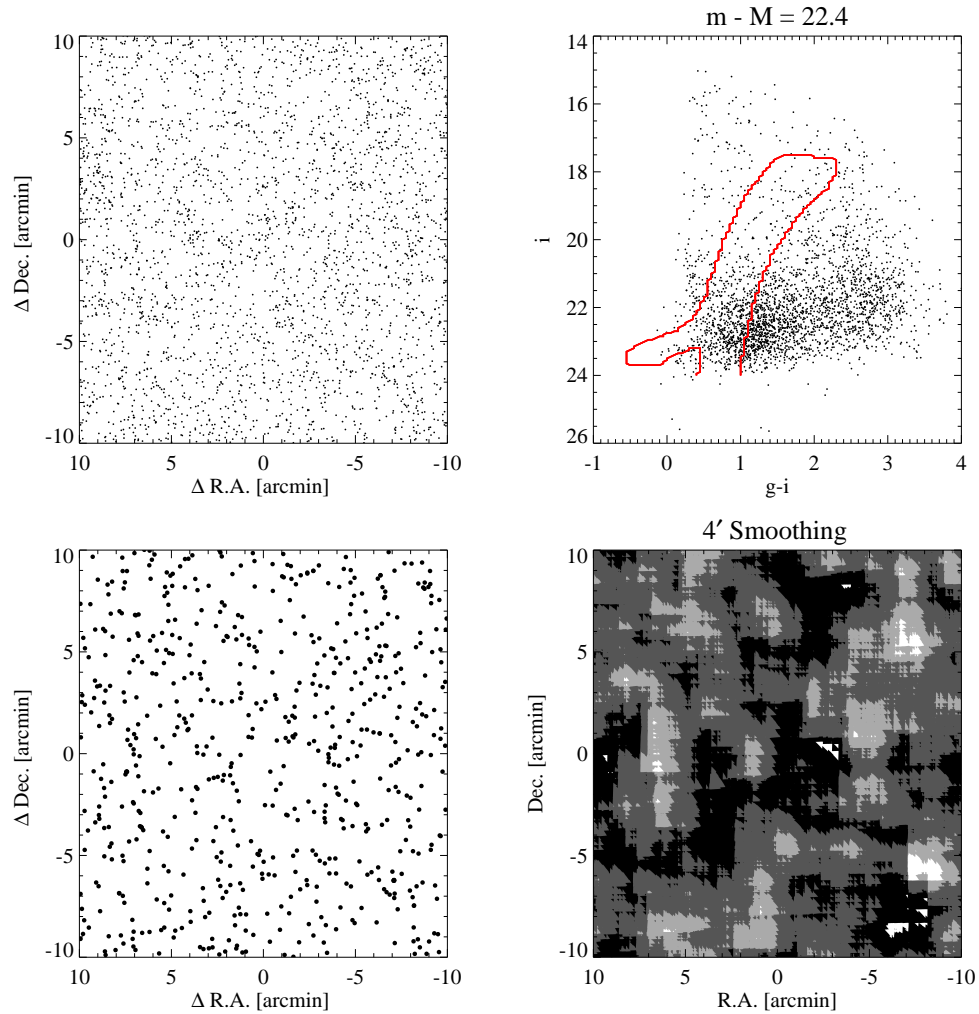


Figure 5.6: An illustration of the source detection methodology used on the field of AGC268069. First, all sources (ideally only stellar sources) within the field are identified (upper left). The sources are then filtered based on the expected colors and magnitudes for an old, metal poor stellar population at a given distance (upper right; $m-M=22.4$ here). The sources remaining after filtering (bottom left) are then smoothed (bottom right; $4'$ spatial scale in this example) to search for an overdensity. Note that no potential counterpart is detected in the example shown here.

the stellar population of Leo P is $\sim 90''$ in extent [Rhode et al., 2013].

The smoothing was performed by gridding the filtered stars onto an array with a resolution of $1.1''$ (10 pixels), smaller than the separation between any two stars. The FFT of this grid was then multiplied by the FFT of a 2D Gaussian with a FWHM of the chosen spatial scale. The inverse of this multiplication was then taken which, by the convolution theorem, is the convolution of the spatial grid of the stars with the Gaussian kernel. Figure 5.6 shows the result of all the steps in the analysis for a single distance modulus of 22.4.

5.6 Discussion

As a demonstration of the technique, Figure 5.7 shows the filtered population of detected objects smoothed to $4'$ for the full range of distance moduli explored. There is an overdensity of sources at approximately $3.5'$ to the south and $6.5'$ to the east of the field center, most prominent for a distance modulus of $m - M = 24.4$. Figure 5.8 shows the distance modulus of 24.4 smoothed on the four different scales considered. The overdensity is strongest for the scale of $1'$ suggesting this is a compact distant source, such as a galaxy cluster. This hypothesis is quickly confirmed by visual inspection which reveals that the identified sources are clearly galaxies and not stars. The excellent image quality of the WIYN data will allow stars to be distinguished from galaxies in most cases based on a shape or FWHM criterion. In addition, for four systems, we have obtained a third filter (r) to be used for helping to discriminate stars from background galaxies based on CMD

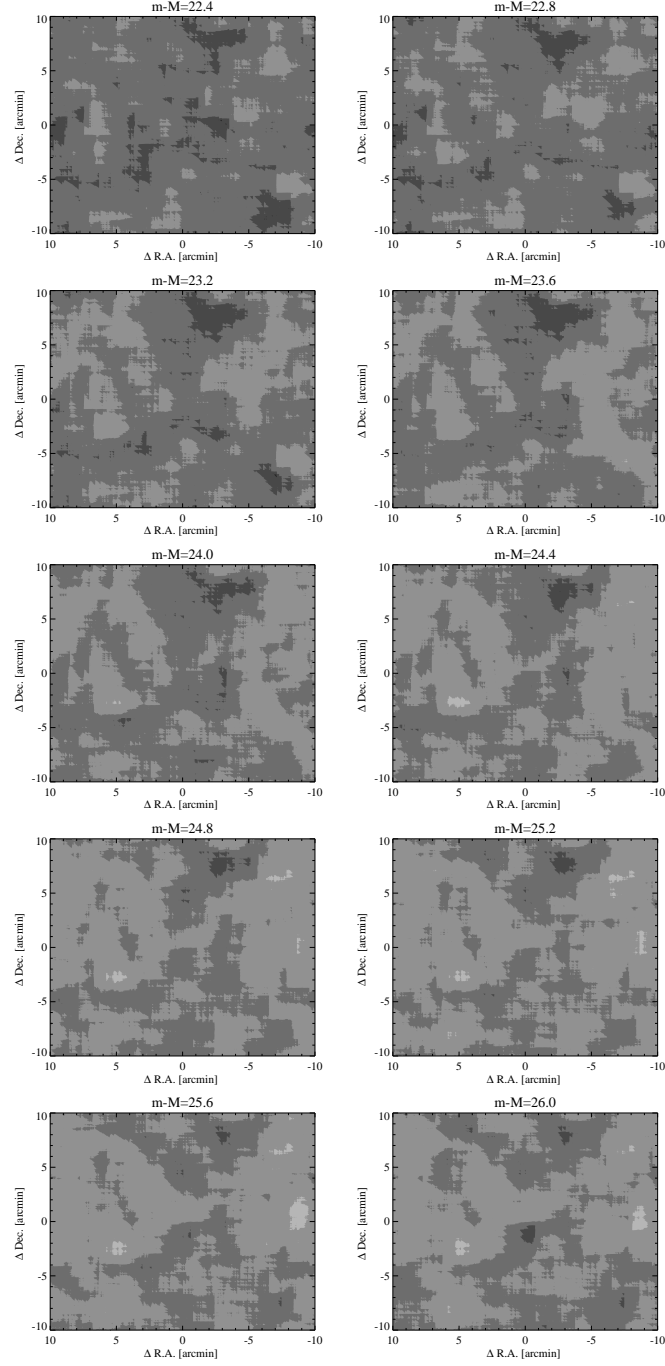


Figure 5.7: The $4'$ smoothed source distribution for all distance moduli considered. Coordinates are given by the offset from the field center. There appears to be an overdensity offset by $3.5'$ to the south and $6.5'$ to the east from the field center. This overdensity is most prominent at $m - M = 24.4$.

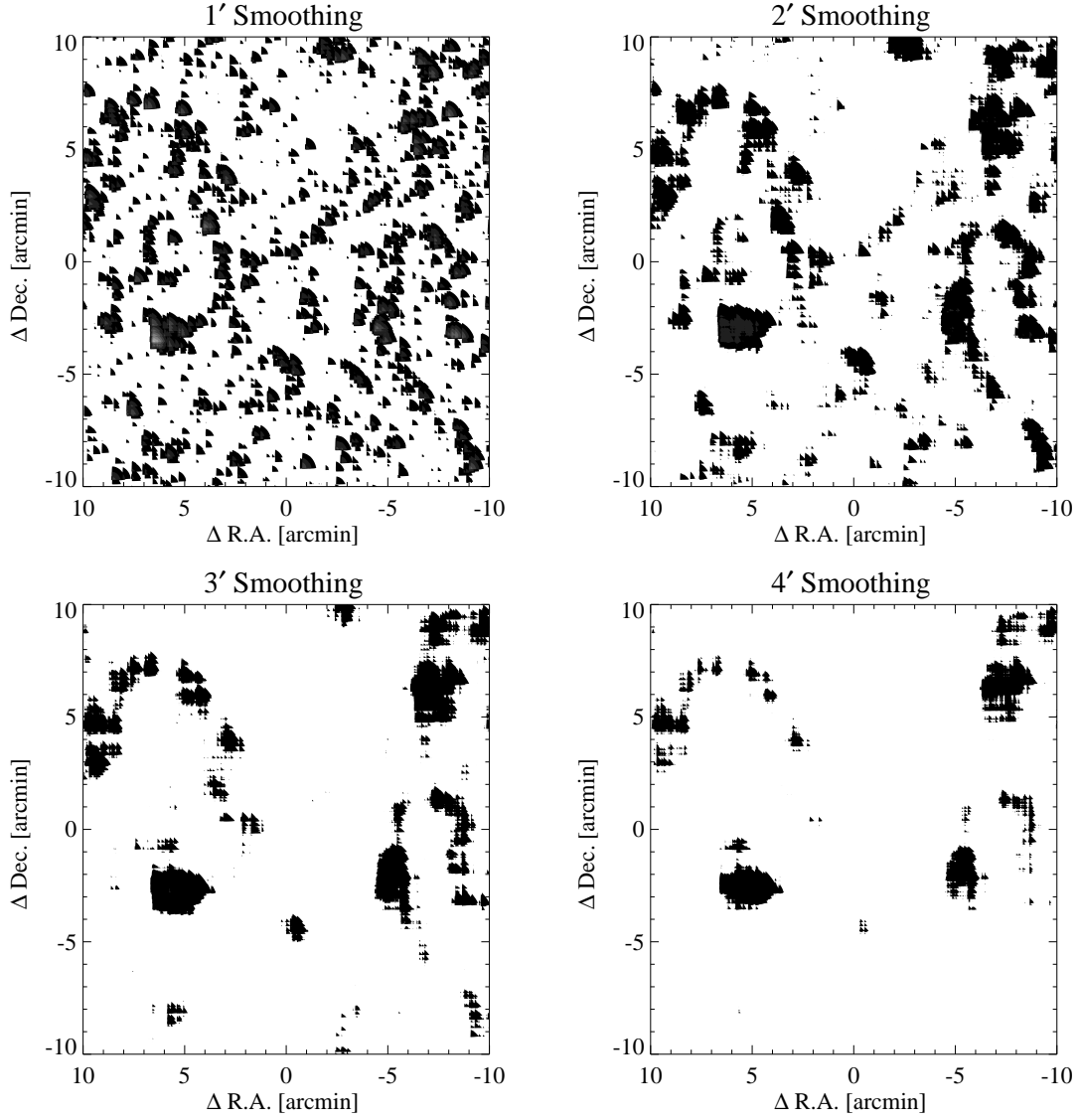


Figure 5.8: Filtered sources for a distance modulus of 24.4 filtered to four different spatial scales. The scaling is the same in all panels. The overdensity is strongest for the smallest scale, suggesting a compact, distant source such as a galaxy cluster. This is verified by visual inspection of the data revealing that all the sources at the location of the overdensity are galaxies.

location.

5.7 Summary

Detecting an optical counterpart to the UCHVCs will definitively determine their nature and constrain their properties via a distance estimate. A best sample of UCHVCs for optical imaging has been selected based on considerations of size, isolation, kinematics and surface brightness. We are undertaking an optical survey of this sample sensitive to the RGB to distances of ~ 1.5 Mpc. Through three different observing programs, we have obtained data for 28 objects in two filters (g and i) and 4 objects in three filters (g, i and r). These data will allow us to perform a filtered search for old, metal-poor stellar populations at a variety of distances. Further work is needed both on data calibration (especially image stacking and cosmic ray removal) and analysis (especially star-galaxy separation) but our initial results suggest that we will be able to detect stellar populations or place significant limits on the stellar populations in the UCHVCs.

CHAPTER 6

CONCLUSION AND FUTURE DIRECTIONS

The search for extremely low mass galaxies is a key way to address the “small scale crisis” in cosmology. The discrepancies between predictions by simulations and observations of low mass galaxies are generally accepted to result from various astrophysical processes, such as reionization and star formation feedback. However, this is an ongoing area of research with many open questions. Is there a minimum mass galaxy that can form? At what mass scales can SF feedback drive density profile changes? Can galaxies with a cored profile survive interaction with the MW? What observed galaxy properties are the result of evolution versus environment?

In this dissertation, ultra-compact high velocity clouds discovered within the ALFALFA HI survey are presented as good candidates to represent low mass gas-rich galaxies within the Local Group. Importantly, if any of these systems do indeed represent galaxies, their gas content argues that they have remained isolated from the MW and M31, implying that any observed properties are intrinsic and the result of quiet evolution. One of the biggest obstacles in understanding the evolution of the lowest mass systems is that it is not clear which properties are intrinsic and which are the result of interaction with a massive galaxy. Are the UFDs so faint because they are intrinsically low mass galaxies or because they have been stripped by the Milky Way? As galaxies, the UCHVCs would suffer less from this ambiguity and allow the underlying evolutionary processes to be probed more directly.

6.1 Identification of UCHVCs

The UCHVCs are identified within the ALFALFA dataset via a customized 3D matched filter signal identification algorithm presented in Chapter 2. Compared to standard extragalactic HI detections within ALFALFA, the UCHVCs are spatially extended and have narrow velocity widths, motivating the development of the specialized algorithm. The efficiency of the algorithm was tested with the insertion of simulated sources, and it outperforms the standard ALFALFA source identification scheme for identifying UCHVCs. In addition, measuring the simulated sources shows that the global properties of the UCHVCs (θ_{HI} , W_{50}) are well measured, but the axial ratios are not remotely accurate.

Chapter 3 presented the catalog of UCHVCs from the $\alpha.40$ footprint. UCHVCs are selected based on a compact size ($\theta_{HI} < 30'$), separation from the Galactic HI ($|v_{LSR}| \geq 120 \text{ km s}^{-1}$) and isolation from surrounding structure. UCHVCs are required to be isolated from large scale HVC structure from the WvW catalog plus small scale structure uncovered in the ALFALFA dataset. In addition, an extra isolation criterion defines a set of most-isolated (MIS) UCHVCs. The isolation tests in Chapter 2 indicate that the MIS UCHVCs are the best minihalo candidates. The full catalog includes 59 UCHVCs and 19 MIS UCHVCs.

6.2 The UCHVCs as a Class

Chapter 3 also examines the properties of the UCHVCs as a class and argues for the minihalo hypothesis. The presence of the Magellanic Stream and other HVC

structure in the fall sky make it unreasonable to argue for any other origins for those HI detections. However, in the spring sky the spatial and kinematic distribution of the UCHVCs is consistent with simulations of dark matter halos within the LG. In addition, the HI properties of the UCHVCs (if placed at 1 Mpc) are consistent with both theoretical and observational predictions for low mass gas-rich galaxies. Importantly, the HI properties of the UCHVCs are consistent with those of two recently discovered low mass gas-rich galaxies in the Local Group and Local Volume, Leo T and Leo P.

We do emphasize that while the galaxy hypothesis is a strong hypothesis and likely true for at least some of the UCHVCs, we do not expect every UCHVC to a gas-bearing dark matter halo. Just like with the general HVC population, we expect the UCHVCs arise from a variety of origins.

6.3 Case Studies of UCHVCs

One way to address the hypothesis that the UCHVCs represent gas in low mass dark matter halos is to study individual UCHVCs in more detail to search for evidence that can discriminate between the opposing cases of either an extragalactic distance (~ 1 Mpc) or a nearby Galactic or circumgalactic origin ($\lesssim 250$ kpc). In Chapters 4 and 5, detailed follow-up observations of specific UCHVCs were presented, in addition to the motivation behind these observations.

6.3.1 High Spatial Resolution HI Studies

High resolution HI studies allow indirect commentary on the minihalo hypothesis. Evidence of interaction of the HI with an ambient medium, such as ram pressure stripping, bow-shock shape, etc., offers strong evidence for a nearby distance for the UCHVCs. Contrarily, a smooth spherical HI appearance is evidence for a distant, undisturbed object. The velocity field of high resolution observations can also be used to address the minihalo hypothesis – does the system show evidence for ordered rotation and velocity support? In that case, there is likely a hosting dark matter halo. Finally, high resolution observation allow detailed HI radial profiles to be derived and offer the resolution to spatially resolve multi-phase cores, allowing close comparisons to models of gas in dark matter halos.

In Chapter 4, we examined high resolution ($\sim 15''$) VLA/C HI observations of AGC268609, a UCHVC with HST-COS observations of a background quasar. The observations revealed the presence of a clumpy HI distribution with at least four major components: three unresolved cores and an irregularly shaped extended northern component. The VLA observations only recovered $\sim 50\%$ of the total flux of AGC268069, indicating that about half the mass is located in an extended envelope. The presence of the northern extension may be extremely tentative evidence for interaction. However, observations of the extended envelope and its morphology are necessary to offer any real constraint on the environment of AGC268069.

6.3.2 Optical Searches for Stellar Counterparts

One direct way to address the minihalo hypothesis is by identification of a stellar counterpart to a UCHVC. The detection of a stellar component would immediately confirm that a UCHVC is a galaxy and would also allow a determination of the distance via the stellar properties.

Chapter 5 presented an ongoing program to search for stellar counterparts to the UCHVCs. The best galaxy candidates from the current survey coverage were selected based on compactness, isolation, high column density and large recession velocities. The systems were observed in two filters as a focused search for a red giant branch population. A preliminary analysis of one target, AGC198606, was undertaken to demonstrate the planned methodology. This first look at the data did not show any evidence of a stellar counterpart but as discussed in Chapter 5 and below in Section 6.4.2, further work with the data is necessary before any definitive conclusion can be drawn on the stellar population.

6.4 Future Work

The galaxy hypothesis for the UCHVCs is very intriguing but requires much more work to fully explore. As shown by the case studies in Chapters 4 and 5, detailed follow-up observations of the UCHVCs can be used to address the galaxy hypothesis but require substantial data. In addition, future deep HI surveys can address the hypothesis by detecting UCHVCs in other galaxy groups.

6.4.1 Expanding the UCHVC Catalog

A clear extension of this work is to produce a catalog of UCHVCs for the final ALFALFA survey coverage. Importantly, expanded survey coverage affects the isolation of sources already identified, especially those near current survey edges. With expanded survey coverage, it will be worthwhile to return to the question of setting isolation criteria for the UCHVCs. In addition to a criterion on the number of neighbors within a given distance, nearest neighbor distances will be explored for the utility of separating UCHVCs from the general HVC population in the ALFALFA dataset. The catalog will also be expanded to velocities including Galactic hydrogen. Saul et al. [2012] developed a method for identifying compact clouds within the GALFA-HI survey that searches Galactic and non-Galactic velocities; this algorithm is being modified to run on other data sets, including the ALFALFA data.

6.4.2 Optical Studies

One definitive way to address the minihalo hypothesis for the UCHVCs is through detection of a stellar counterpart. In addition to our ongoing work imaging selected UCHVCs, future large optical surveys will be able to detect these systems.

Targeted Optical Studies of UCHVCs

As outlined in Chapter 5, there are 25 UCHVCs with observations in 2 broadband filters (g and i), and 4 with observations in 3 broadband filters (g, r, and i). We

hope to detect stellar counterparts in at least one of these systems and to detect as many as 5 would be significant. The non-detections will also provide informative upper limits on the stellar population in the UCHVCs.

Chapter 5 presented a preliminary reduction and analysis of one target but refinement is needed to the methodology presented there. Importantly, the ODI pipeline is under development and analysis either needs to await an improved version of the pipeline (anticipated fall 2013) or more interaction is required to reduce the pODI data. Currently, fringe corrections are not included in the pipeline and are necessary for the *i* filter. In addition, the pipeline stacking does not deal properly with the background and the individual OTAs need to be combined by the user.

The identification of stellar sources can also be improved, with the key improvement being star-galaxy discrimination. The excellent image quality of the WIYN pODI data will allow this to be done via constraints on the FWHM. In addition, four fields were observed in 3 filters so multiple CMD cuts can be applied to help remove background galaxies. Identification of stellar sources will also be improved by more accurate handling of the background, allowing for fainter sources to be robustly detected. In order to place accurate limits on stellar populations, the detection limits of the images must be well understood. This will be done by adding artificial stellar sources to the images and running them through the analysis steps.

Future Optical Surveys

Future optical surveys will also expand the search for UFDs to distances of 1 Mpc or more. The Dark Energy Survey (DES) and PanSTARRS will be sensitive to UFDs with total luminosities of a $\sim \text{few} \times 10^4 L_{\odot}$ out to distances of ~ 1 Mpc while LSST will be able to detect systems with total luminosities of a $\text{few} \times 10^3 L_{\odot}$ to those distances [Tollerud et al., 2008]. In addition, there are targeted wide-field surveys of nearby galaxies, such as NGC 253 and NGC 5128, to search for UFDs in other systems [Sand, 2013].

6.4.3 High Resolution HI Imaging

Chapter 4 addresses how high spatial resolution HI observations can constrain the minihalo hypothesis for the UCHVCs. Deep observations of AGC268069 with the VLA in C-configuration were presented as a case study; similar observations exist for a second system (AGC258242) that also has HST-COS data of a background quasar. Together these datasets with the combination of absorption spectra and high resolution HI imaging will allow stringent comparisons of the two UCHVCs to models of gas in low mass dark matter halos.

However, the data for AGC268069 does need further work, most especially to robustly test and deal with the elevated flux levels. Further data cubes will be created with a longer spectral baseline (to higher recessional velocities), and cleaning to different levels and over different sized boxes will be implemented to address the pedestaling issue. Alternative continuum source subtraction strategies will also be tested to see if the elevated flux issue lies instead in poor continuum sub-

traction. Even after tweaking the calibration and imaging of this dataset, $\sim 50\%$ of the flux will still be missing from the observations. A proposal has been placed to observe this source, AGC258242, and eight other UCHVCs that have optical imaging with the VLA in D-configuration. As the most compact configuration, the D array has the best column density sensitivity and does the best job of recovering the flux on extended scales. These observations will be crucial to constraining the environment of AGC268069 and other UCHVCs through the large scale morphology of the envelope.

An ongoing program of 169 hours will map 7 UCHVCs with the Westerbork Synthesis Radio Telescope (WSRT), providing spatial resolution of $30''$. This program will produce a consistent and uniform dataset that can be used to examine the HI morphology of UCHVCs as a class, searching for evidence of interaction with the circumgalactic medium of the MW indicating that they are at distances < 300 kpc or evidence that the UCHVCs are rotation-supported structures indicating that they represent gas in low mass dark matter halos.

6.4.4 Future HI Surveys

One possibility is that the UCHVCs represent low mass dark matter halos which have only recently (re)acquired neutral gas as the Universe expands and cools [e.g., Ricotti, 2009]. In this scenario, there is a possibility that the UCHVCs do truly represent dark matter halos but have no detectable stellar counterpart to constrain the distance. In this case, optical observations will not reveal a stellar counterpart, making it near impossible to directly prove the minihalo hypothesis as the HI detection carries no direct distance information. One way around this

is to undertake deep surveys of nearby (3-5 Mpc) galaxy groups. The spatial and kinematic association of an HI detection with a galaxy group allows the group distance estimate to be confidently applied to the HI detection. As discussed in Chapter 3, extant HI surveys of nearby galaxy groups lack either the field of view or depth necessary to systematically constrain the presence of $10^5 M_{\odot}$ HI clouds at distances of ~ 1 Mpc from the massive central galaxies.

A new generation of instruments under development, focal-plane phased-array feeds, will provide the large field of views necessary for such surveys. In general these instruments are being developed as pathfinders for the Square Kilometre Array (SKA) but are being deployed on current telescopes or pathfinder arrays for the SKA and will enable deep HI surveys in advance of the full SKA. One such instrument is Apertif, which is currently being deployed at ASTRON for WSRT. Apertif is a focal-plane array that will increase the field of view of WSRT by a factor of 25. This increased field of view makes the mapping speed for deep surveys reasonable, and planned projects for Apertif include deep HI surveys to search for $10^5 M_{\odot}$ objects out to several Mpc.

6.5 Summary

The UCHVCs are good candidates to represent low mass galaxies within the Local Group. As galaxies, the UCHVCs would be among the lowest mass gas-rich objects known and would place important constraints on baryonic processes in low mass dark matter halos. Exploring the minihalo hypothesis for the UCHVCs is challenging but ongoing work to search for stellar counterparts and examine

the HI morphology and kinematics of the UCHVCs will offer insight to the hypothesis.

APPENDIX A

SOURCE CODE FOR THE 3D EXTRACTION ALGORITHM FOR UCHVCS

In this appendix, I present the 3D signal extraction code used to identify the UCHVCs within the ALFALFA data set. The subprocedures called by the main code `uchvc_ex3d` are presented first, followed by the main code.

```

;=====
; uchvc_ex3d: 3D extraction of UCHVC/minihalo candidates
; Written by BA to identify UCHVCs within the
; ALFALFA dataset
;=====

;=====
function twod_4degpoly,x,y,p
; indep variables x & y are sigma_spat & sigma_vel,
; respectively
; p is an array that contains all the coeffiecients for
; the various polynomial degrees

; this is a 4th deg poly because higher order is too
; complicated!

zmod=p[0]+p[1]*x+p[2]*y+p[3]*x*y+p[4]*x^2+$
      p[5]*y^2+p[6]*x^2*y+p[7]*x*y^2+$
      p[8]*x^3+p[9]*y^3+p[10]*x^3*y+$
      p[11]*x^2*y^2+p[12]*x*y^3+p[13]*x^4+p[14]*y^4

return ,zmod
end
;=====

;=====
pro create_cube ,grid ,cube ,weights ,pol0 ,pol1
; create a data cube truncated to the dimensions needed
; for UCHVC search

common uchvcvals ,ampl ,sz_spat ,sz_vel ,N ,sig_sp ,sig_v , $
      mx,my,mz ,peakflux ,intflux ,cmin ,cmax ,ramin ,ramax , $
      decmin ,decmax ,sn ,noise ,color_pol ,pol_flag

```

```

;manually set cube dimensions here
cube=reform( (grid.d[435:734,0,*,*]+$
              grid.d[435:734,1,*,*])/2. )
weights=reform( (grid.w[435:734,0,*,*]+$
                 grid.w[435:734,1,*,*])/2. )
pol0=reform( grid.d[435:734,0,*,*])
pol1=reform( grid.d[435:734,1,*,*])

;return sz_spat , sz_vel
sz_spat=n_elements(cube[0,*,0])
sz_vel=n_elements(cube[*,0,0])

end
;=====

;=====
pro find_galhi ,cube ,galind ,nogalind ,slice
;identify velocity slices where Galactic HI is present

common uchvcvals ,ampl ,sz_spat ,sz_vel ,N ,sig_sp ,sig_v , $
        mx,my,mz ,peakflux ,intflux ,cmin ,cmax ,ramin ,ramax , $
        decmin ,decmax ,sn ,noise ,color_pol ,pol_flag

slice=fltarr(sz_vel)
for i=0,sz_vel-1 do slice[i]=total(abs(cube[i,*,*]))
meanslice=mean(slice)
medianslice=median(slice)
cutoff=medianslice+0.15*medianslice
nogalind=where(slice lt cutoff)
galind=where(slice ge cutoff)

end
;=====

;=====
pro remove_galhi ,cubein ,galind ,cubeout
;blank the Galactic HI

common uchvcvals ,ampl ,sz_spat ,sz_vel ,N ,sig_sp ,sig_v , $
        mx,my,mz ,peakflux ,intflux ,cmin ,cmax ,ramin ,ramax , $
        decmin ,decmax ,sn ,noise ,color_pol ,pol_flag

```

```

cubeout=cubein
cubeout[galind,*,*]=0.

end
;=====

;=====
pro check_galhi,cube,galreg
;check the a reasonable part of the data cube is
;identified as Galactic HI

common uchvcvals,ampl,sz_spat,sz_vel,N,sig_sp,sig_v,$
        mx,my,mz,peakflux,intflux,cmin,cmax,ramin,ramax,$
        decmin,decmax,sn,noise,color_pol,pol_flag

find_galhi,cube,galind

;find out how many regions there are and what they are

ngalreg=1
galreg = [galind[0]]
nelemgal=n_elements(galind)
for i=0,nelemgal-2 do begin
    if galind[i+1]-galind[i] gt 1 then begin
        galreg=[galreg,galind[i],galind[i+1]]
        ngalreg=ngalreg+1
    endif
endfor
if (galind[nelemgal-1] ne galreg[n_elements(galreg)-1]) $
then galreg=[galreg,galind[nelemgal-1]]

print, 'Number of Galactic HI regions is ', $
        strtrim(string(ngalreg),2)
print, 'They are: '
for i=0,ngalreg-1 do print, $
        galreg[i*2]+435, ': ', galreg[i*2+1]+435

end
;=====

;=====

```

```

pro calc_globalnoise ,cube ,nogalind ,globalnoise
; calculate global noise of cube in Galaxy-free channels

common uchvcvals ,ampl ,sz_spat ,sz_vel ,N ,sig_sp ,sig_v , $
      mx,my,mz,peakflux ,intflux ,cmin ,cmax ,ramin ,ramax , $
      decmin ,decmax ,sn ,noise ,color_pol ,pol_flag

nogalHI=cube (nogalind )
size_nogalHI=n_elements (nogalHI)
globalnoise=sqrt ( total (nogalHI^2) / size_nogalHI)

end
;=====

;=====
pro find_source ,cube ,fft_d ,template ,narr ,sources ,cnt , $
      ptsflag

common uchvcvals ,ampl ,sz_spat ,sz_vel ,N ,sig_sp ,sig_v , $
      mx,my,mz,peakflux ,intflux ,cmin ,cmax ,ramin ,ramax , $
      decmin ,decmax ,sn ,noise ,color_pol ,pol_flag

; take FFTed data and templates and do the convolution ,
; find convolution maximum for each template
nt=n_elements (template )
cnvmax=fltarr (nt)
cx=fltarr (nt)
cy=fltarr (nt)
cz=fltarr (nt)
cnv=fltarr (nt ,sz_vel ,sz_spat ,sz_spat)
; loop through all the templates and find location of peak
; have to manipulate index for each dimension
for i=0,nt-1 do begin
      cnv [i ,* ,* ,* ]=$
          real_part (FFT (( fft_d *conj (template [i] . fft_t )) , $
                          /INVERSE) / sqrt (N))
      max=max (cnv [i ,* ,* ,* ] ,mind) ; look for a maximum
      cnvmax [i ]=max
      cz [i ]=mind mod sz_vel
      cx [i ]= (mind /sz_vel) mod sz_spat
      cy [i ]= mind / (sz_vel *sz_spat)
endfor

```

```

;first set the location of the detection
a=max(cnvmax,aind)
mx=cx[ aind ]
my=cy[ aind ]
mz=cz[ aind ]
tmp_s = template[ aind ].sigma_spat
tmp_v = template[ aind ].sigma_vel

;make sure that the list of convolution values
;I input to the fitting function
;are all taken from the same location
cnvvals=cnv[* ,mz,mx,my]

;reset the cnv array to zero to free up memory
cnv=0.

;now we want to find the parameters of the
;best matched template by
;interpolating between the set of fixed template values
;set up arrays that cover the range of the sigma values
maxsigs=max(template.sigma_spat)
minsigs=min(template.sigma_spat)
maxsigv=max(template.sigma_vel)
minsigv=min(template.sigma_vel)

sigs=dblarr(1001,1001)
sigv=dblarr(1001,1001)

for i=0,1000 do sigs[i,*]=minsigs+$
    i/1000.*(maxsigs-minsigs)
for i=0,1000 do sigv[* ,i]=minsigv+$
    i/1000.*(maxsigv-minsigv)

;now calculate the fit cmax values
err=fltarr(nt)+0.01
start_params=fltarr(15)+1.0

parms=mpfit2dfun('twod_4degpoly',template.sigma_spat,$
    template.sigma_vel,cnvvals,err,$
    start_params,/quiet)

```



```

alpha=twod_4degpoly(sigs ,sigv ,parms)

;the good cnv and sigma values
alpha_max=max(alpha ,ind)
sig_sp=sigs(ind)
sig_v=sigv(ind)
ampl=alpha_max*sqrt(N)/$
      sqrt(sqrt(!dpi^3)*sig_sp*sig_sp*sig_v)
;and now I have my best template values
;while i could use the calculated peak value
;i'll go ahead and
;go through the calculation of peak flux
;and such below to ensure the
;most accuracy since that's the peak
;calculated from the data

;set limits of detection based on sigmas of detection
cmin = floor(mz-2*sig_v)
if cmin lt 0 then cmin=0
cmax = ceil(mz+2*sig_v)
if cmax gt (sz_vel-1) then cmax=sz_vel-1
ramin = floor(mx-2*sig_sp)
if ramin lt 0 then ramin=0
ramax=ceil(mx+2*sig_sp)
if ramax gt (sz_spat-1) then ramax=sz_spat-1
decmin=floor(my-2*sig_sp)
if decmin lt 0 then decmin=0
decmax = ceil(my+2*sig_sp)
if decmax gt (sz_spat-1) then decmax=sz_spat-1

;isolate the source and calculate peak and total flux
source=cube[cmin:cmax,ramin:ramax,decmin:decmax]
peak=max(source)
indm=where(source gt 0.75*peak)
peakflux=total(source[indm])/n_elements(indm)
intflux=5.1*total(total(total(source,3),2),1)

;check number of points in detection
ptsflag=0
if n_elements(indm) le 5 then ptsflag=1

print,'The source has spatial and velocity sizes of:',$

```

```

2.3548*sig_sp,2.3458*5.1*sig_v

print,'The peak and fitted amplitude are',peakflux,ampl

end
;=====

;=====
function make_source

common uchvcvals,ampl,sz_spat,sz_vel,N,sig_sp,sig_v,$
      mx,my,mz,peakflux,intflux,cmin,cmax,ramin,ramax,$
      decmin,decmax,sn,noise,color_pol,pol_flag

a=max([ampl,peakflux])

width = 2.3548*5.1*sig_v

sc=a*$
      psf_gaussian(npixel=[sz_vel,sz_spat,sz_spat],$
                    fwhm=[2.3548*sig_v,2.3548*sig_sp,$
                          2.3548*sig_sp],$
                    centroid=[mz,mx,my],$
                    ndimension=3,/double)

return,sc

end
;=====

;=====
pro good_sn,cube,weights
;do a good S/N calculation based on Amelie's code

common uchvcvals,ampl,sz_spat,sz_vel,N,sig_sp,sig_v,$
      mx,my,mz,peakflux,intflux,cmin,cmax,ramin,ramax,$
      decmin,decmax,sn,noise,color_pol,pol_flag

;do a good SN calc here ignoring contribution of source
w1=weights[cmin:cmax,0:ramin,my]
w2=weights[cmin:cmax,ramax:sz_spat-1,my]
sw1=n_elements(w1)
sw2=n_elements(w2)

```

```

wrms=fltarr (sw1+sw2)
wrms[0:sw1-1]=w1
wrms[sw1:(sw1+sw2-1)]=w2

d1=cube[cmin:cmax,0:ramin,my]
d2=cube[cmin:cmax,ramax:sz_spat-1,my]
sd1=n_elements(d1)
sd2=n_elements(d2)
datarms=fltarr(sd1+sd2)
datarms[0:sd1-1]=d1
datarms[sd1:(sd1+sd2-1)]=d2

nodataps=where(wrms gt 0.0)
if (n_elements(nodataps) eq 1) then begin
    if (nodataps lt 0) then goto,skipthenextline2
endif
datarms=datarms[nodataps]
skipthenextline2:
chanrms=sqrt(total(datarms^2)/n_elements(datarms))

within3sig=where(abs(datarms) lt 3*chanrms,nbrgood)
chanrms=sqrt(total(datarms[within3sig]^2)/nbrgood)

datarms=datarms[within3sig]
within3sig=where(abs(datarms) lt 3*chanrms,nbrgood)
chanrms=sqrt(total(datarms[within3sig]^2)/nbrgood)

noise=chanrms
sn=peakflux/noise

w=sig_v*5.1*2.3568
if (w lt 20.) then sn=sn
if (w ge 20.) then sn=sn*sqrt(w/(2*10.))
;don't have to worry about the w>200 correction to S/N
;as I don't expect sources that large

print, 'Noise, peak flux, and SN values are: ', $
    noise, peakflux, sn

end
;=====

```

```

;=====
pro check_pol , pol0 , pol1

common uchvcvals , ampl , sz_spat , sz_vel , N , sig_sp , sig_v , $
      mx , my , mz , peakflux , intflux , cmin , cmax , ramin , ramax , $
      decmin , decmax , sn , noise , color_pol , pol_flag

      stt = noise / (sqrt(cmax-cmin+1)*sqrt(decmax-decmin+1)*$
                    sqrt(ramax-ramin+1))

      p0=total(pol0[cmin:cmax, ramin:ramax, decmin:decmax])/$
        (sqrt(cmax-cmin+1)*sqrt(decmax-decmin+1)*$
        sqrt(ramax-ramin+1))
      p1=total(pol1[cmin:cmax, ramin:ramax, decmin:decmax])/$
        (sqrt(cmax-cmin+1)*sqrt(decmax-decmin+1)*$
        sqrt(ramax-ramin+1))

      sav=0.5*abs(p0+p1)
      diff_pol=abs(p0-p1)/stt

      if (sav lt 50*stt) then begin
        if (diff_pol gt (20.)) then begin
; polarisations are obviously different
          pol_flag=1
          color_pol='0000FF'XL
        endif
        if (diff_pol gt (8.) and diff_pol le (20.))$
        then begin
; marginal difference between pols
          pol_flag=7
          color_pol='00FF00'XL
        endif
        if (diff_pol le (8.)) then begin
; good detection
          color_pol='00FF00'XL
          pol_flag=0
        endif
      endif else begin
        if (diff_pol gt (0.3*sav/stt)) then begin
          pol_flag=1
          color_pol='0000FF'XL
        endif

```

```

        if (diff_pol gt (0.1*sav/stt) and $
diff_pol le (0.3*sav/stt)) then begin
            pol_flag=7
            color_pol='00FF00'XL
        endif
        if (diff_pol le (0.1*sav/stt)) then begin
            color_pol='00FF00'XL
            pol_flag=0
        endif
    endelse

end
;=====

;=====
pro ex3d,grid,sources,res,snth=snth

;define a common block for use in all the programs
;to carry valuess around
common uchvcvals,ampl,sz_spat,sz_vel,N,sig_sp,sig_v,$
    mx,my,mz,peakflux,intflux,cmin,cmax,ramin,ramax,$
    decmin,decmax,sn,noise,color_pol,pol_flag

;check sn threshold
if n_elements(snth) eq 0 then snth=4.3

res=fltarr(1024,144,144)

;trim data to cube
create_cube,grid,cube,weights,pol0,pol1

;find and remove Galactic HI
print,'Finding and blanking the Galactic hydrogen'
find_galhi,cube,gallnd,nogallnd
tmp_cube=cube
remove_galhi,tmp_cube,gallnd,cube
tmp_cube=0.

;restore the templates to extract sources
restore,'/home/humacao/humacao2/betsey/chvcextractor/$
idl/template_expanded.sav'

```

```

;start recording the system time
t0=systime(1)

;create the structure that sources will be stored in
;will truncate extra entries at end
entry={ch:0,ra:0,dec:0,w:0.0,ara:0.0,adec:0.0,sn:0.0,$
        int_flux:0.0,peak_flux:0.0,rms:0.0,$
        diff_pol:0.0,pol_flag:0,color_pol:'',cz:0.0,$
        agc:'',comments:''}
sources=replicate(entry,100)

;set initial values before beginning loop
s=15.
sn=10.
cnt=0
peak=10.
snloop=15.
cnt_nd=0
N=n_elements(cube)
nt=n_elements(template)
loopn=0

print, 'Beginning loop to find sources'
while (snloop ge snth) do begin
    loopn=loopn+1
    print, 'On loop number ',loopn

    print, 'Taking the FFT of the data'
    fft_d=FFT(cube)*sqrt(N)

    ;do the cnv and find first source
    print, 'Looking for a cnv peak'
    find_source,cube,fft_d,template,narr,sources,cnt,ptsflag

;do a good SN calc here ignoring contribution of source
    print, 'Doing the good SN calc'
    good_sn,cube,weights

;set the value in res array
    res[mz+435,mx,my]=sn

    print, 'the source is located at ',mz,mx,my

```

```

;run a check on polarization levels
;this code comes straight from amelie's extractor
  print,'Checking pols'
  check_pol,pol0,pol1

  print,'Subtracting the matched source'
  ;we subtract from the cube for calculating
;source properties

  src=make_source()
  cube=cube-src

;reset source to zero
  src=0.

;check that we're above the threshold value and
;add sources to catalog
  if (sn ge snth) then begin
    print,'Updating source catalog'
    ;update sources
    ;am not calculating all the params yet
    ;so i don't update all of them
    sources[cnt].ch=mz+435
;channel number in original grid
    sources[cnt].ra=round(mx)
    sources[cnt].dec=round(my)
    sources[cnt].w=2.3568*sig_v*5.1
    sources[cnt].ara=2.3568*sig_sp
    sources[cnt].adec=2.3568*sig_sp
    sources[cnt].peak_flux=peakflux
    sources[cnt].int_flux=intflux
    sources[cnt].sn=sn
    sources[cnt].cz=grid.velarr[mz+435]
    sources[cnt].pol_flag=pol_flag
    sources[cnt].color_pol=color_pol
    if (sources[cnt].pol_flag eq 7) then $
      sources[cnt].comments='pol.diff.'
    ;update the count
    cnt=cnt+1
  endif

```

```

        snloop=sn

endwhile

t1=systime(1)
print , 'It took ',strtrim( string((t1-t0)/60.),2),$
        ' minutes to find the sources'

;truncate the sources catalog
if cnt eq 0 then begin
    print , 'No sources detected ', $
        ' creating an empty file for ',grid.name
    sources=sources[0:cnt]
endif else begin
    sources=sources[0:cnt-1]
;write out to a file
    print , 'Saving sources for ',grid.name
    print , 'There are ',cnt,' sources'
    save ,sources ,res ,filename='chvc_'+grid.name+'.sav'
endelse

end

```


APPENDIX B

DETAILED REDUCTION AND IMAGING STEPS FOR VLA DATA

This appendix contains the full details on the calibration and imaging of the VLA dataset discussed in Chapter 4. All calibration and imaging was done in CASA unless otherwise noted.

B.1 Calibration

Measured visibilities consist of two quantities - an amplitude and a phase. The amplitude contains information about the strength of the signal and the phase contains information about the location of emission for the source observed. In order to calibrate the observed visibilities of the source, calibrators are observed. A flux calibrator is a source of known intensity; it allows the measured amplitudes to be converted to a flux scale. A bandpass calibrator (typically the flux calibrator for low frequency observations) is a strong source that is used to empirically determine the response of the observing set-up as a function of frequency. The phase calibrator is a point source located near the target used to calibrate the phases. Observations of the phase calibrator occur through the same atmosphere as the source and since it is a point source its expected phases are zero, allowing the offsets in the phases of each antenna to be determined.

B.1.1 Flagging

Before calibrating the data it is necessary to identify and flag bad data. The on-line observing log is first checked to identify any bad antennas that need to be removed. Then the `viewer` within CASA is used to examine the data interactively, starting with the flux calibrator and looking at baselines as a function of time. Bad baselines or antennas are identified and note for later flagging. Then the phase calibrator and source are also checked. Next the data is examined for RFI by inspecting channels as function of time for given baseline pairs. First a few baseline pairs for the calibrators and source are checked. If there is strong RFI present, all baselines and polarizations (for all SPWs) need to be examined and flagged interactively using the flagging tools within the `viewer`.

once finished examining data with the `viewer`, `plotms` is used to examine the dataset. Before viewing the data in `plotms`, the current flags are applied to the data to enable the best autoscaling in `plotms` with the following command:

```
tflagdata(vis=vis5,mode='list',infile='flag_5.list')
```

This reads all the flags in the list and applies them; importantly it only requires the visibility measurement set to be loaded a single time for all flags to be applied, saving considerable computational time. A flag list may look like the following:

```
antenna='4'  
antenna='22' correlation='LL'  
antenna='26' timerange='05:55:46~05:56:09'
```

In `plotms`, the amplitudes are examined as a function of time for each field, iterating by baseline, averaging all channels together and colorizing by SPW or

polarization. Any bad baselines, antennas or timeranges are noted for flagging. Then the phase is examined as a function of time for the flux and phase calibrators. The phase should vary smoothly for the calibrators and have a small range of values. Occasionally, the phases will vary wildly if there is a problem, such a loose cable, in the system somewhere. Then the affected antenna will need to be flagged. Any new flags are applied before the last check for RFI.

The presence of any RFI is hopefully minor, or taken care of previously in the viewer. The presence of any minor RFI is checked for by plotting amplitude as function of frequency and averaging over time (start by averaging scans separately, then combine the scans). and iterating by baseline. Any channels affected by RFI are noted and flagged. A list of flag commands may look like the following:

```
antenna='ea09&ea16' scan='15~18' spw='3:53~58,11:117~122'
antenna='ea09&ea16' spw='3:55,11:119'
antenna='ea09&ea22' spw='3:55,11:119'
antenna='ea09&ea22' scan='6~15' spw='3:54~56,11:118~119'
antenna='ea16&ea19' spw='3:55,11:119'
```

B.1.2 Split the Data

Before proceeding with the data calibration, we split the dataset by polarization. This is because CASA will not calibrate an antenna if a single polarization is flagged and it expects there to be two polarizations. Splitting the data so that only one polarization is expected provides a work around. The commands for splitting the data and defining the new visibilities for further work are:

```
split(vis=vis5,outputvis='sb9400530_LL.ms',
      correlation='LL',datacolumn='data')
```

```
split(vis=vis5,outputvis='sb9400530_RR.ms',
      correlation='RR',datacolumn='data')

visLL5='sb9400530_LL.ms'
visRR5='sb9400530_RR.ms'
```

B.1.3 Set the Flux Scale

We use stored models of the flux calibrators to determine the flux within our observed frequency range of the flux calibrator and set the flux scale:

```
setjy(vis=visLL5,field='0',modimage='3C286_L.im')
setjy(vis=visRR5,field='0',modimage='3C286_L.im')
```

B.1.4 Bandpass Calibration

The next step is to calibrate the bandpass response of the observations. First, the phases are calibrated with time since the bandpass scans (flux calibrator) are fairly long (~ 8 minutes) and are separated by the full length of the observing block. We pick a few central channels to average over for the phase calibration so that there is enough signal but bandpass effects are not important, since we have not done that correction yet. The antenna 'ea02' has been chosen as a reference antenna as it appears well-behaved and is near the center of the array (use `plotants` to determine antenna location), providing a good range of baselines. Calibrate the phases:

```
gaincal(vis=visLL5,caltable='bpphaseLL.gcal',
        field='0',spw='*:55~75',
```

```

        refant='ea02', calmode='p', solint='60s',
        minsnr=2.0, opacity=0.0, gaincurve=T)
gaincal(vis=visRR5, caltable='bpphaseRR.gcal',
        field='0', spw='*:55~75',
        refant='ea02', calmode='p', solint='60s',
        minsnr=2.0, opacity=0.0, gaincurve=T)

```

In general, we wish to examine our solutions to confirm that they look reasonable. This can be done with the task `plotcal`:

```

plotcal(caltable='bpphaseLL.gcal', xaxis='time',
        yaxis='phase', iteration='antenna', subplot=331)
plotcal(caltable='bpphaseRR.gcal', xaxis='time',
        yaxis='phase', iteration='antenna', subplot=331)

```

The phases look reasonable so proceed to doing the bandpass calibration, applying the phase solutions derived above:

```

bandpass(vis=visLL5, caltable='bandpassLL.bcal', field='0',
        refant='ea02', solint='inf', solnorm=T,
        gaintable=['bpphaseLL.gcal'],
        opacity=0.0, gaincurve=T)
bandpass(vis=visRR5, caltable='bandpassRR.bcal', field='0',
        refant='ea02', solint='inf', solnorm=T,
        gaintable=['bpphaseRR.gcal'],
        opacity=0.0, gaincurve=T)

```

Once again, check the solutions. This time we want to examine both the phase and amplitude solutions. In addition, we can choose to show the solutions as function of frequency or channel (will overlap the SPWs).

```

plotcal(caltable='bandpassLL.bcal', yaxis='phase',
        xaxis='freq', iteration='antenna', subplot=331)
plotcal(caltable='bandpassLL.bcal', yaxis='amp',

```

```

        xaxis='freq', iteration='antenna', subplot=331)
plotcal(caltable='bandpassRR.bcal', yaxis='phase',
        xaxis='chan', iteration='antenna', subplot=331)
plotcal(caltable='bandpassRR.bcal', yaxis='amp',
        xaxis='chan', iteration='antenna', subplot=331)

```

Note that the channels 30–100 have a near uniform response in all SPWs; these channels will be the ones used for the upcoming calibration steps.

B.1.5 Gain Calibration

We are now ready to calibrate the complex gains of the calibrators. First we will calibrate the phases on a 60 second time scale within a scan. The phases should be relatively constant for low frequency L-band observations but this ensures that we are accounting for variability within a scan. These phase solutions will be used when deriving the amplitude solution for the calibrators. Calculate the phase solutions, applying the bandpass calibration:

```

gaincal(vis=visLL5, caltable='intphaseLL.gcal',
        field='0,1', spw='*:30~100',
        refant='ea02', calmode='p', solint='60s',
        minsnr=2.0, gaintable=['bandpassLL.bcal'],
        opacity=0.0, gaincurve=T)
gaincal(vis=visRR5, caltable='intphaseRR.gcal',
        field='0,1', spw='*:30~100',
        refant='ea02', calmode='p', solint='60s',
        minsnr=2.0, gaintable=['bandpassRR.bcal'],
        opacity=0.0, gaincurve=T)

```

And check the solutions:

```

plotcal(caltable='intphaseLL.gcal', yaxis='phase',

```

```

        xaxis='time', iteration='antenna', subplot=331)
plotcal(caltable='intphaseRR.gcal', yaxis='phase',
        xaxis='time', iteration='antenna', subplot=331)

```

We now wish to derive a single phase value for the entire scan for use when applying to the source. It is helpful to record the number of possible solutions and the number of good solutions to ensure that the data is well calibrated.

```

gaincal(vis=visLL5, caltable='scanphaseLL.gcal',
        field='0,1', spw='*:30~100',
        refant='ea02', calmode='p', solint='inf',
        minsnr=2.0, gaintable=['bandpassLL.bcal'],
        opacity=0.0, gaincurve=T)
#256 solution intervals; 224 good solutions
gaincal(vis=visRR5, caltable='scanphaseRR.gcal',
        field='0,1', spw='*:30~100',
        refant='ea02', calmode='p', solint='inf',
        minsnr=2.0, gaintable=['bandpassRR.bcal'],
        opacity=0.0, gaincurve=T)
#256 solution intervals; 224 good solutions

```

As always, check the solutions. Then calibrate the amplitude of the calibrators over each scan, using the 60 second phase solutions from above:

```

gaincal(vis=visLL5, caltable='ampLL.gcal',
        field='0,1', spw='*:30~100',
        refant='ea02', calmode='ap', solint='inf',
        minsnr=2.0, opacity=0.0, gaincurve=T,
        gaintable=['bandpassLL.bcal', 'intphaseLL.gcal'])
gaincal(vis=visRR5, caltable='ampRR.gcal',
        field='0,1', spw='*:30~100',
        refant='ea02', calmode='ap', solint='inf',
        minsnr=2.0, opacity=0.0, gaincurve=T,
        gaintable=['bandpassRR.bcal', 'intphaseRR.gcal'])

```

Check the solutions, noting that the phases should be zero as they have been calibrated out at this point in time.

B.1.6 Get the Flux Scale

We now wish to get the flux for the phase calibrator based on the known flux of the flux calibrator. Essentially we are calibrating the amplitudes to an absolute value; this calibration will be transferred to the source later, setting the flux scale of the source also. Use the task `fluxscale` for this:

```
fluxscale(vis=visLL5, caltable='ampLL.gcal',  
          fluxtable='fluxLL.cal', reference='0')  
fluxscale(vis=visLL5, caltable='ampRR.gcal',  
          fluxtable='fluxRR.cal', reference='0')
```

A list of flux values (one for each SPW) should print to the terminal. Record these values and compare them to the listed flux value of the source in the VLA Calibrator Manual. The values may vary somewhat as the phase calibrators are variable sources, but they should be relatively the same. An example of the calculated fluxes for the phase calibrator use here, J1553+1256, are:

```
LL: ([ 1.16919878, 1.16840626, 1.16915217, 1.1692162,  
       1.17332324, 1.15152723, 1.16899301, 1.16832557,  
       1.16557518, 1.16657125, 1.16733276, 1.1661959 ,  
       1.16522234, 1.16834083, 1.16171913, 1.16739391])  
  
RR: ([ 1.16983337, 1.17257224, 1.17309703, 1.17124434,  
       1.17113017, 1.15158464, 1.1722049, 1.17125246,  
       1.1713119, 1.17129231, 1.170885, 1.17104463,  
       1.17112099, 1.1717945, 1.1671281, 1.17334829])
```

These values compare well to the value in the manual of 1.10 Jy.

The calculated amplitudes can also be check with `plotcal`:

⁰<http://www.vla.nrao.edu/astro/calib/manual/>


```

plotcal(caltable='fluxLL.cal',yaxis='amp',xaxis='time',
        iteration='antenna',subplot=331)
plotcal(caltable='fluxRR.cal',yaxis='amp',xaxis='time',
        iteration='antenna',subplot=331)

```

B.1.7 Apply the Calibration

We are now ready to apply the solutions to the source and calibrators. We will apply the solutions to the sources one a time, paying special attention to which gain tables are relevant to which source, and which source (`gainfield`) solutions should be used. The bandpass solutions always use the flux calibrator (`field='0'`) as it is the only source for which a bandpass correction is calculated. For the flux calibrator, the reference field is always itself. For the phase calibrator and source, the reference field is the phase calibrator, except for the bandpass correction. Note that the calibrators use the phase solutions on the 60 second timescale while the source uses the scan averaged phase solutions.

```

applycal(vis=visLL5,field='0',
         gaintable=['bandpassLL.bcal','intphaseLL.gcal',
                   'fluxLL.cal'],
         gainfield=['0','0','0'],
         opacity=0.0,gaincurve=T,calwt=F)
applycal(vis=visLL5,field='1',
         gaintable=['bandpassLL.bcal','intphaseLL.gcal',
                   'fluxLL.cal'],
         gainfield=['0','1','1'],
         opacity=0.0,gaincurve=T,calwt=F)
applycal(vis=visLL5,field='2',
         gaintable=['bandpassLL.bcal','scanphaseLL.gcal',
                   'fluxLL.cal'],
         gainfield=['0','1','1'],
         opacity=0.0,gaincurve=T,calwt=F)
applycal(vis=visRR5,field='0',

```

```

    gaintable=['bandpassRR.bcal','intphaseRR.gcal',
               'fluxRR.cal'],
    gainfield=['0','0','0'],
    opacity=0.0,gaincurve=T,calwt=F)
applycal(vis=visRR5,field='1',
    gaintable=['bandpassRR.bcal','intphaseRR.gcal',
               'fluxRR.cal'],
    gainfield=['0','1','1'],
    opacity=0.0,gaincurve=T,calwt=F)
applycal(vis=visRR5,field='2',
    gaintable=['bandpassRR.bcal','scanphaseRR.gcal',
               'fluxRR.cal'],
    gainfield=['0','1','1'],
    opacity=0.0,gaincurve=T,calwt=F)

```

B.1.8 Check Calibration and Redo as Necessary

At this stage we wish to check our calibration one last time. If we notice any bad antennas or baselines, we can flag them and redo the steps above. Typically, the easiest thing to do is to create new calibration tables that are named with `_redo` after flagging the data, and then to apply those new tables to the (flagged) source data.

Many calibration issues will be noticed during the steps above when examining the solutions with `plotcal`. At this stage, the best way to examine the data for calibration issues is to plot the calibrated phases and amplitudes against each other in `plotms`. As the calibrator sources are point sources, the phases and amplitudes should occupy a smooth circular distribution centered at a phase of 0 and an amplitude that is the flux of the source. When displaying the data it is important to average either in time or spectrally (or both) in order to see the points clustered enough that outliers are evident. As with previous work in `plotms`, iter-

ating by baseline and/or antenna and coloring by various properties (antenna2, scan, SPW) will help identify systemically bad antennas/baselines/scans.

B.1.9 Split and “cvel” the Data

We want to split out just the calibrated source for further work. This produces a smaller dataset and allows us to easily return to the current visibilities if a later problem is noticed with data calibration (e.g. striping when imaging). Split the calibrated source data into a new file:

```
split(vis=visLL5, field='2', datacolumn='corrected',  
      outputvis='sourceLL.ms')  
split(vis=visRR5, field='2', datacolumn='corrected',  
      outputvis='sourceRR.ms')
```

We now wish to `cvel` the data. As discussed in §4.4.2, the data are observed in overlapping spectral bands and we wish to create a single dataset with uniform spectral coverage. This can be done by selecting the channels from all SPWs to produce uniform full spectral coverage and then “gluing” them together. In this case, we use `cvel` to do this as the channels used in this observational setup are narrow enough that signal shifts from one channel to another during the full scheduling block due to the Earth’s rotation; `cvel` accounts for this shift.

```
cvel(vis='sourceLL.ms',  
     spw='0~6:31~94,7:31~127,8:0~94,9~15:31~94',  
     restfreq='1420.40575MHz', veltype='optical',  
     outframe='bary',  
     mode='frequency', outputvis='cvel_sourceLL.ms')  
cvel(vis='sourceRR.ms',  
     spw='0~6:31~94,7:31~127,8:0~94,9~15:31~94',
```

```
restfreq='1420.40575MHz',veltype='optical',
outframe='bary',
mode='frequency',outputvis='cvel_sourceRR.ms')
```

B.1.10 Continuum Subtraction

Finally, we wish to subtract the continuum sources before proceeding to imaging. The left panel of Figure B.1 shows a dirty image of the continuum sources in the field that need to be subtracted. First we make dirty cubes that can be used to identify line-free (both source and Galactic HI in this case) channels for use in the continuum subtraction:

```
clean(vis='cvel_sourceLL.ms',imagename='dirty_cvel_LL',
      niter=0,restfreq='1420.40575MHz',mode='frequency',
      outframe='bary',cell='5arcsec',imsize=512)
clean(vis='cvel_sourceLL.ms',imagename='dirty_cvel_RR',
      niter=0,restfreq='1420.40575MHz',mode='frequency',
      outframe='bary',cell='5arcsec',imsize=512)
```

We examine the resulting images in `viewer` to determine that the channels for continuum subtraction are: [30, 360] and [1025, 1060]. Then do the continuum subtraction:

```
uvcontsub(vis='cvel_sourceLL.ms',
          fitspw='0:30~360;1025~1060')
uvcontsub(vis='cvel_sourceRR.ms',
          fitspw='0:30~360;1025~1060')
```

We make new dirty cubes to verify that the continuum subtraction worked; examine the cubes in `viewer` after they are produced.

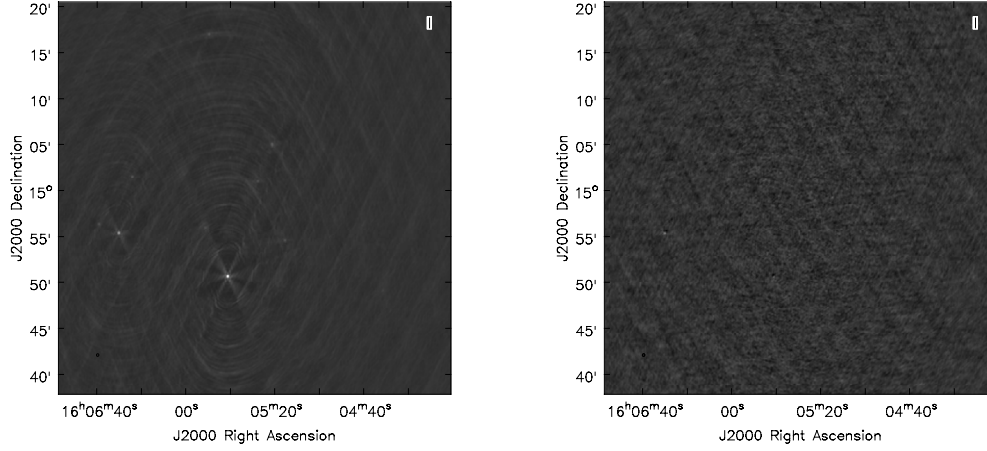


Figure B.1: The dirty continuum image of a single scheduling block before subtraction of continuum sources (left) and after continuum source subtraction (right). Both images are shown to the same gray scale.

```
clean(vis='cvel_sourceLL.ms.contsub',
      imagename='dirty_cvel_LL_line',
      niter=0, restfreq='1420.40575MHz', mode='frequency',
      outframe='bary', cell='5arcsec', imsize=256)
clean(vis='cvel_sourceLL.ms.contsub',
      imagename='dirty_cvel_RR_line',
      niter=0, restfreq='1420.40575MHz', mode='frequency',
      outframe='bary', cell='5arcsec', imsize=256)
```

The right panel of Figure B.1 shows a continuum image of the field after the continuum source subtraction, demonstrating that the continuum subtraction is effective. However, there is some evidence for residuals left from the continuum subtraction; the effects of this are discussed briefly in §B.2.3. Future work will include testing different subtraction methods and perhaps smoothing spectrally in the uv plane before subtraction.

B.2 Imaging

This section describes in detail the procedures used to create the final moment maps and spectra presented in Chapter 4. The creation of moment maps is, itself, straightforward. The challenging parts include creating a good mask and accounting for pedestaling in the data cubes which artificially elevates the flux level.

B.2.1 Obtaining Clean Image Cubes

Before creating cleaned data cubes, dirty cubes are created for all imaging schemes and frequency resolutions. These cubes are used to obtain the cleaning level. As an example, the naturally-weighted 8 kHz resolution dirty cube is created as follows:

```
clean(vis=[source1LL, source1RR, source2LL, source2RR,
          source3LL, source3RR, source4LL, source4RR,
          source5LL, source5RR],
      imagename='dirty_natural_8kHz',
      niter=0, restfreq='1420.40575MHz', mode='frequency',
      start='1419.503165MHz', nchan=64,
      outframe='bary', cell='5arcsec', imsize=512,
      width='7.9125kHz', weighting='natural',
      uvtaper=False)
```

The noise level of the cube in the inner box where cleaning will occur is retrieved automatically and used to set the threshold for clean:

```
xstat=imstat(imagename='dirty_natural_8kHz.image',
             chans='2', verbose=F, box='204,204,304,304')
```

```

rms_nat_8 = xstat['rms'][0]
thres_nat_8 = str(rms_nat_8) + 'Jy'
clean(vis=[source1LL, source1RR, source2LL, source2RR,
          source3LL, source3RR, source4LL, source4RR,
          source5LL, source5RR],
      imagename='clean_natural_8kHz',
      niter=10000, restfreq='1420.40575MHz',
      mode='frequency', outframe='bary', cell='5arcsec',
      imsize=512, width='7.9125kHz', weighting='natural',
      uvtaper=False, nchan=64, start='1419.503165MHz',
      mask=[204, 204, 304, 304], multiscale=[0, 4, 16],
      threshold=thres_nat_8)

```

B.2.2 Creating a Mask

In order to produce final spectra and moment maps, we wish to produce a mask that can be used for all data cubes to only allow the region with real signal to be considered but also allowing relevant noise to contribute. The tapered cube at 16 kHz resolution was determined to be the best cube for this through examination.

First, the cube was smoothed to about twice its resolution ($60'' \times 60''$) and clipped at the 2 sigma level. Note that to be truly accurate the clipping should occur after accounting for the elevated flux level (discussed in §B.2.3); this will be done in future work to create a better mask.

```

imsmooth(imagename='clean_taper_16kHz.image',
        kernel='gauss', major='60arcsec',
        minor='60arcsec', targetres=True,
        outfile='smooth_16kHz.im')
xstat=imstat(imagename='smooth_16kHz.im', chans='1',
            verbose=F, box='204, 204, 304, 304')
rms_16 = xstat['rms'][0]

print 2*rms_16

```

```
0.0037373355589807034
```

```
ia.open('smooth_16kHz.im')
ia.calcmask(mask='smooth_16kHz.im >0.003737',
            name='clipmask1')
ia.done()
```

Then the clipped image was interactively masked using the masking tool in imager toolkit (used for creating interactive masks when cleaning). In order to be considered “real” emission, signal needed to be significant and connected contiguously to signal in an adjacent channel. The channels that were interactively selected to have signal were: [14, 21] (or [120.4, 143.7] km s⁻¹). The masking tool actually produces an image, so it was turned into a CASA mask and a master mask was created.

```
im.drawmask(image='smooth_16kHz.im', mask='blank16.mask')
#page through the channels and highlight the signal
#when all emission is selected, quit the interaction and
#save mask by hitting the blue arrow
im.close()
```

```
ia.open('blank16.mask')
ia.calcmask(' "blank16.mask">0.5')
ia.close()
```

```
#use immath to create master masked iamges
immath(imagename='smooth_16kHz.im', mode='evalexpr',
        expr='IM0', outfile='master16.im',
        mask='mask(blank16.mask)')
```

The mask created here (varying in shape for each channel) can be used to produce a very clean moment map. However, we wish to create a mask that is uniform in frequency so that noise is being properly included in final spectra and moment maps. We will apply the mask we have just created to the unsmoothed data,

create a moment map, and use the produced moment map (potentially clipped) as the final mask, propagating it to all channels of the data cube. Before applying the mask to the unsmoothed data, we create a copy of the data to apply the mask to. `maskhandler` should be able to handle any issues with masks, including masking and unmasking at will. However, logistically it is easier to track images if the original data cubes are kept unmasked. CASA requires a mask to have the same shape as the image it is applied to, so the mask produced by the moment zero map will need to be expanded in frequency space for the three different resolution cubes. There is no straight-forward way to do this in CASA, so we export the moment zero image and tapered data cubes (for size of the velocity axis) to fits file and use IDL code provided by G. Hallenbeck to create 3D masks.

```
#copy the image
immath(imagename='clean_taper_16kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='temp_taper_16kHz.im')

#apply the mask
ia.open('temp_taper_16kHz.im')
ia.maskhandler('copy',['master16.im:mask0','newmask'])
ia.maskhandler('set','newmask')
ia.done()

#make a moment zero map
#channel range doesn't matter due to mask
immoments(imagename='temp_taper_16kHz.im',moments=[0],
           axis='spectral',mask='',chans='',
           outfile='moment0_mask_16kHz')

#export the files
exportfits(imagename='moment0_mask_16kHz',
           fitsimage='mom0_mask_16kHz.fits')
#also need to export a data cube that will be used to
#propagate mask to the proper dimensions
exportfits(imagename='clean_taper_8kHz.image',
           fitsimage='8kHz_taper_cube.fits')
```

```

exportfits(imagename='clean_taper_16kHz.image',
           fitsimage='16kHz_taper_cube.fits')
exportfits(imagename='clean_taper_32kHz.image',
           fitsimage='32kHz_taper_cube.fits')

#in idl:
#.comp mask
#mask,'8kHz_taper_cube.fits','mom0_mask_16kHz.fits',$
#'mask_8kHz.fits',cutoff=0.0001
#mask,'16kHz_taper_cube.fits','mom0_mask_16kHz.fits',$
#'mask_16kHz.fits',cutoff=0.0001
#mask,'32kHz_taper_cube.fits','mom0_mask_16kHz.fits',$
#'mask_32kHz.fits',cutoff=0.0001

#now import the masks into CASA:
importfits(fitsimage='mask_8kHz.fits',
           imagename='master_mask_8kHz')
importfits(fitsimage='mask_16kHz.fits',
           imagename='master_mask_16kHz')
importfits(fitsimage='mask_32kHz.fits',
           imagename='master_mask_32kHz')

#turn the master mask images into actual masks
ia.open('master_mask_8kHz')
ia.calcmask(' "master_mask_8kHz">0.5')
ia.summary()
ia.close()
ia.open('master_mask_16kHz')
ia.calcmask(' "master_mask_16kHz">0.5')
ia.summary()
ia.close()
ia.open('master_mask_32kHz')
ia.calcmask(' "master_mask_32kHz">0.5')
ia.summary()
ia.close()

#copy data to new images and mask it at same time
immath(imagename='clean_taper_8kHz.image',
       mode='evalexpr',expr='IM0',
       outfile='mask_taper_8kHz.im',
       mask='mask(master_mask_8kHz)')
immath(imagename='clean_taper_16kHz.image',
       mode='evalexpr',expr='IM0',

```

```

        outfile='mask_taper_16kHz.im',
        mask='mask(master_mask_16kHz)')
immath(imagename='clean_taper_32kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='mask_taper_32kHz.im',
        mask='mask(master_mask_32kHz)')
immath(imagename='clean_natural_8kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='mask_natural_8kHz.im',
        mask='mask(master_mask_8kHz)')
immath(imagename='clean_natural_16kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='mask_natural_16kHz.im',
        mask='mask(master_mask_16kHz)')
immath(imagename='clean_natural_32kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='mask_natural_32kHz.im',
        mask='mask(master_mask_32kHz)')
immath(imagename='clean_robust_8kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='mask_robust_8kHz.im',
        mask='mask(master_mask_8kHz)')
immath(imagename='clean_robust_16kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='mask_robust_16kHz.im',
        mask='mask(master_mask_16kHz)')
immath(imagename='clean_robust_32kHz.image',
        mode='evalexpr',expr='IM0',
        outfile='mask_robust_32kHz.im',
        mask='mask(master_mask_32kHz)')

```

B.2.3 Pedestal Effect

As seen in Figure 4.5 in Chapter 4, the initially produced spectra using the masked data cubes created above clearly show that the spectra are not baselined properly. Most likely, this is the result of the deep cleaning (to one sigma) resulting in a “pedestal” to the flux level. Alternatively, it could be the result of a poorly sub-

tracted continuum source and future work will involve revisiting the continuum subtraction. The elevation of the pedestal was calculated for the cubes made with different imaging schemes using channels [0, 7] and [26, 31] in the 16kHz cubes as signal-free channels. The calculated pedestal levels for the three different cubes are:

Tapered cube pedestal level: 0.000299139Jy/bm

Natural cube pedestal level: 0.000179078Jy/bm

Robust cube pedestal level: 2.91040e-05Jy/bm

These values were subtracted from the masked data cubes to create the data cubes used for further analysis:

```
immath(imagename='mask_taper_16kHz.im',mode='evalexpr',
       expr='IM0-0.000299139',
       outfile='mask_baselined_taper_16kHz.im')
immath(imagename='mask_natural_16kHz.im',mode='evalexpr',
       expr='IM0-0.000179078',
       outfile='mask_baselined_natural_16kHz.im')
immath(imagename='mask_robust_16kHz.im',mode='evalexpr',
       expr='IM0-2.91040e-05',
       outfile='mask_baselined_robust_16kHz.im')
```

B.2.4 Final Products: Moment Maps and Spectra

Spectra were created from the data cubes using `imstat` to record the sum in each channel. For each differently imaged cube, the conversion to flux (Jy) from the sum (Jy/bm) was different as the number of beams contained within the mask varied as a result of the different beam sizes. An example of how `imstat` was called to get the spectra of a data cube:

```
imstat(imagename='mask_baselined_taper_16kHz.im',
```

```
axes=[0,1,2],
logfile='imstat_mask_baselined_taper_16kHz_freq.log')
```

After creating the spectra, moment zero (integrated flux density) maps are created for the 16kHz cubes for all three imaging schemes. The cubes were created for two different velocity ranges: the larger channel range encompassing all signal seen in the spectra and a more restricted velocity range encompassing the channels for which signal was selected during the masking step above:

```
immoments(imagename='mask_baselined_taper_16kHz.im',
           moments=[0],axis='spectral',chans='11~23',
           outfile='mom0_baselined_taper')
immoments(imagename='mask_baselined_taper_16kHz.im',
           moments=[0],axis='spectral',chans='14~21',
           outfile='mom0_baselined_taper_restricted')
```

The moment zero maps were converted to column density using the formula:

$$T_B = \frac{606 S_{21}}{a \times b} \quad (\text{B.1})$$

$$N_{HI} = 1.823 \times 10^{18} \int T_B dv \quad (\text{B.2})$$

$$(\text{B.3})$$

where S_{21} is the flux density in mJy, and a and b are the beam sizes. Explicitly this done in CASA as:

```
immath(imagename='mom0_baselined_taper',mode='evalexpr',
       expr='1.823e18 * 606*1e3 / (26.5*25.3) * IM0',
       outfile='mom0_baselined_nhi_taper')
immath(imagename='mom0_baselined_taper_restricted',
       mode='evalexpr',
       expr='1.823e18 * 606*1e3 / (26.5*25.3) * IM0',
       outfile='mom0_baselined_nhi_taper_restricted')
```

A moment one map, or velocity field, was created for the tapered data cube over the limited velocity range. The moment zero map clipped at the 4×10^{19} atoms cm^{-2} level was used a mask to isolated the most significant emission.

```
#create the velocity field
immoments(imagename='mask_baselined_taper_16kHz.im',
           moments=[1],axis='spectral',chans='14~21',
           outfile='velfield_baselined_restricted')

#copy moment zero map to be used as a mask
immath(imagename='mom0_baselined_nhi_taper_restricted',
       mode='evalexpr',expr='IM0',
       outfile='mask_velfield_restricted')

#create the mask
ia.open('mask_velfield_restricted')
ia.calcmask(mask='mask_velfield_restricted > 4e19',
            name='clipmask')
ia.done()

#copy the mask to the velocity field
ia.open('velfield_baselined_restricted')
ia.maskhandler('copy',
               ['mask_velfield_restricted:clipmask',
                'newmask'])
ia.maskhandler('set','newmask')
ia.done()
```

B.3 Image Testing

Several test were performed to determine which imaging strategy to use for the VLA data. Here, details on running some of those tests are discussed.

B.3.1 Combining SPWs

In order to determine the best way to combine the overlapping SPWs, we imaged a single scheduling block in three different ways: using all channels, using only channels with good bandpass performance and by “gluing” channels together to create a single bandpass.

First, “glue” the channels together:

```
cvel(vis='sourceLL.ms',
     spw='0~6:31~94,7:31~127,8:0~94,9~15:31~94',
     restfreq='1420.40575MHz',veltype='optical',
     outframe='bary',
     mode='frequency',outputvis='cvel_sourceLL.ms')
cvel(vis='sourceRR.ms',
     spw='0~6:31~94,7:31~127,8:0~94,9~15:31~94',
     restfreq='1420.40575MHz',veltype='optical',
     outframe='bary',
     mode='frequency',outputvis='cvel_sourceRR.ms')
```

Then image the original data with all channels and only the central channels of each SPW, plus the glued data created above:

```
clean(vis=[sourceLL,sourceRR],
     imagename='dirty_allfreq',
     niter=0,restfreq='1420.40575MHz',mode='frequency',
     outframe='bary',cell='5arcsec',imsize=512)
clean(vis=[sourceLL,sourceRR],
     imagename='dirty_selfreq',
     spw='*:30~100',
     niter=0,restfreq='1420.40575MHz',mode='frequency',
     outframe='bary',cell='5arcsec',imsize=512)
clean(vis=[cvel_sourceLL,cvel_sourceRR],
     imagename='dirty_cvel',
     niter=0,restfreq='1420.40575MHz',mode='frequency',
     outframe='bary',cell='5arcsec',imsize=512)
```

Finally, examine the noise as a function of channel for all the dirty images produced above. Output the results to a text file that can be used to plot the results, shown in Figure 4.2 in Chapter 4.

```
imstat(imagename='dirty_good_allfreq.image',
       axes=[0,1,2],logfile='imstat_good_allfreq.log')
imstat(imagename='dirty_good_selffreq.image',
       axes=[0,1,2],logfile='imstat_good_selffreq.log')
imstat(imagename='dirty_good_cvel.image',
       axes=[0,1,2],logfile='imstat_good_cvel.log')
imstat(imagename='dirty_bad_allfreq.image',
       axes=[0,1,2],logfile='imstat_bad_allfreq.log')
imstat(imagename='dirty_bad_selffreq.image',
       axes=[0,1,2],logfile='imstat_bad_selffreq.log')
imstat(imagename='dirty_bad_cvel.image',
       axes=[0,1,2],logfile='imstat_bad_cvel.log')
```

B.3.2 Robustness Levels

In order to test which robustness level to image the final data cubes at (Section 4.5.1), we produced dirty images for a single scheduling block for a variety of robustness levels (-2 to 2, incremented by 0.1), plus pure natural and uniform weighting. The dirty images were produced for only 300 channels to save computational time. The visibilities are first defined globally and then the various robust cubes are produced in a loop. The natural and uniform cubes are produced outside the loop as special cases.

```
msLL = 'sb9400530/cvel_sourceLL.ms.contsub'
msRR = 'sb9400530/cvel_sourceRR.ms.contsub'
for i in range(0,41):
    #set the robustness value
    #will do natural and uniform weighting separately
```



```

rob=i/10. -2.
#first make the dirty image
dirty='dirty_rob'+str(rob)
clean(vis=[msLL,msRR],imagename=dirty,
      niter=0,restfreq='1420.40575MHz',
      mode='frequency',
      outframe='bary',cell='5arcsec',imsize=256,
      weighting='briggs',robust=rob,spw='0:200~555')

dirty='dirty_natural'
clean(vis=[msLL,msRR],imagename=dirty,
      niter=0,restfreq='1420.40575MHz',
      mode='frequency',
      outframe='bary',cell='5arcsec',imsize=256,
      weighting='natural',spw='0:200~555')
dirty='dirty_uniform'
clean(vis=[msLL,msRR],imagename=dirty,
      niter=0,restfreq='1420.40575MHz',
      mode='frequency',
      outframe='bary',cell='5arcsec',imsize=256,
      weighting='uniform',spw='0:200~555')

```

After producing the dirty cubes, retrieve the noise and beam sizes and print to screen. Copy the output to a text file and use it for producing a figure to choose the best robustness values (Figure 4.4 in Chapter 4):

```

for i in range (0,41):
    rob=i/10.-2.
    dirty='dirty_rob'+str(rob)
    dirtyimage=dirty+'.image'
    xstat=imstat(imagename=dirtyimage,verbose=F)
    rms=xstat['rms'][0]
    bmaj = imhead(imagename=dirtyimage,mode='get',
                  hdkey='beammajor')
    bmin = imhead(imagename=dirtyimage,mode='get',
                  hdkey='beamminor')
    print rob,rms,bmaj['value'],bmin['value']

```

```

dirty='dirty_natural'
dirtyimage=dirty+'.image'
xstat=imstat(imagename=dirtyimage,verbose=F)
rms=xstat['rms'][0]
bmaj = imhead(imagename=dirtyimage,mode='get',
              hdkey='beammajor')
bmin = imhead(imagename=dirtyimage,mode='get',
              hdkey='beamminor')
print dirty,rms,bmaj['value'],bmin['value']

dirty='dirty_uniform'
dirtyimage=dirty+'.image'
xstat=imstat(imagename=dirtyimage,verbose=F)
rms=xstat['rms'][0]
bmaj = imhead(imagename=dirtyimage,mode='get',
              hdkey='beammajor')
bmin = imhead(imagename=dirtyimage,mode='get',
              hdkey='beamminor')
print dirty,rms,bmaj['value'],bmin['value']

```

APPENDIX C
DETAILED REDUCTION AND ANALYSIS STEPS FOR WIYN/PODI
OPTICAL DATA

In this appendix, a detailed log of the reduction and analysis is presented for the pODI data for AGC198606. Chapter 5 provides the context for this Appendix. Typically, reduction occurred in IRAF (with a few minor exceptions) and analysis was started in IRAF (star finding and magnitudes) and then switched to IDL (filtering and smoothing). In addition, I point out places where it may be appropriate to explore alternate methodologies.

The starting data products used for the work here are the pipeline-produced pointing stack images, combined exposures based on a common pointing center (and filter). As discussed in Chapter 5, these data are not yet of final science quality. Specifically, the background scaling is not handled properly in the stacking (D. Harbeck, private communication 2013), fringe corrections are not applied, and there is no cosmic ray removal. In order to have science ready data, it will be necessary to either perform the early data reduction steps (e.g. fringe correction and image combining) manually, or to wait for further development of the pipeline. The pipeline is actively being developed and by the end of 2013 should produce much higher quality data products (K. Rhode, private communication 2013). The data reduction and analysis steps outlined here serve as a guide for the work to be undertaken when final science data are available.

C.1 Post-pipeline Processing

Given the dither pattern, the full $24' \times 24'$ field of view does not have uniform coverage. The images are trimmed to the central $20'$ to focus on the area with the best coverage and for ease of use (computing memory).

```
ecl> imcopy pstack_agc198606_g.fits[3000:14000,3000:14000]
      trim_198606_g
ecl> imcopy pstack_agc198606_i.fits[3000:14000,3000:14000]
      trim_198606_i
```

It should be noted that these image sizes are 11000×11000 pixels, which is larger than IRAF can display; the maximum display size is 8192×8192 .

C.1.1 Cosmic Ray Removal

Before cleaning the image by hand, an automated cosmic ray removal was done using the `cosmicray` task in IRAF. First, `imstat` was used to get a sense of the background.

```
ecl> epar imstat

PACKAGE = imutil
TASK = imstatistics
images = trim* List of input images
(fields = image,midpt) Fields to be printed
(lower = 0.) Lower limit for pixel values
(upper = 60000.) Upper limit for pixel values
(nclip = 0) Number of clipping iterations
(lsigma = 3.) Lower side clipping factor
(usigma = 3.) Upper side clipping factor
(binwidth = 0.1) Bin width of histogram in sigma
```

```
(format   =          yes) Format output and print labels
(cache    =          no) Cache image in memory ?
(mode     =          ql)
```

This outputs the following information to the screen:

```
          IMAGE      MIDPT
trim_198606_g.fits  245.
trim_198606_i.fits  622.4
```

We can use the midpoint values reported as approximate sky values for the images and then calculate the expected variance in sky counts via:

$$\sigma = \frac{\sqrt{SB * G + RN^2}}{G} \quad (C.1)$$

where SB is the sky background, RN the read noise and G the gain of the CCD detector. Since we are dealing with stacks created by averaging nine exposures, the read noise and gain need to be modified to account for the change in noise statistics. The gain (1.36) is multiplied by 9 for an effective value of 12.24, and the read noise (6.52) is multiplied by a factor of $\sqrt{9}$ for an effective value of 19.56. With a sky background of 245, the variance for the g band image is 4.75, and with a sky background of 622, the variance for the i band image is 7.31. We wish to set a threshold of 4 times the variance, and so the thresholds for the g and i filters are, respectively, 19 and 29.

Then we edit the parameters of the cosmicray task and run it. Unfortunately running this task interactively results in a segmentation violation so there was no adjustment or training from the entered parameters.

```
ecl> imred
ecl> crutil
```

```

ecl> epar cosmic
PACKAGE = crutil
TASK = cosmicrays
input   = trim_198606_g  List of images in which to
                        detect cosmic rays
output  = cr_198606_g    List of cosmic ray replaced
                        output images (optional)
(crmasks=                ) List of bad pixel masks
                        (optional)
(thresho=                19.) Detection threshold above mean
(fluxrat=                4.) Flux ratio threshold (in percent)
(npasses=                5) Number of detection passes
(window =                5) Size of detection window
(interac=                no) Examine parameters interactively?
(train  =                no) Use training objects?
(objects=                ) Cursor list of training objects
(savefil=                ) File to save train objects
(plotfil=                ) Plot file
(graphic=                stdgraph) Interactive graphics output
                        device
(cursor =                ) Graphics cursor input
answer  =                no  Review parameters for image?
(mode   =                ql)

```

This task removed a few cosmic rays from the image but the vast majority remained. Given this poor performance, and the difficulty of cleaning images by hand (discussed below), an alternative method is preferable. One option is try the `lacosmic` package from van Dokkum [2001]. This package is available for IRAF, IDL and python. It utilizes a Laplacian cosmic ray identification. Alternatively, if the individual images are stacked manually it would be possible to tweak the combination to be an average with some sort of rejection criterion to aid in the removal of cosmic rays, satellite trails, bright star residuals and other image artifacts.

The majority of the image cleaning was done by hand using `imedit`. This task allows a user to interactively select regions and replace the data located there. This task was called in the following manner:

```

PACKAGE = tv
  TASK = imedit
input    =   cr_198606_i  Images to be edited
output   =   e_198606_i  Output images
(cursor  =                ) Cursor input
(logfile=                ) Logfile of cursor commands
(display=                yes) Display images?
(autodis=                no) Automatic image display?
(autosur=                no) Automatic surface plots?
(apertur=   circular) Aperture type
(radius  =                3.) Substitution radius
(search  =                2.) Search radius
(minvalu=                INDEF) Minimum value to modify
(maxvalu=                INDEF) Maximum value to modify
(buffer  =                1.) Background buffer width
(width   =                2.) Background width
(xorder  =                2) Background x order
(yorder  =                2) Background y order
(value   =                0.) Constant value substitution
(sigma   =                INDEF) Added noise sigma
(angh    =               -33.) Horizontal viewing angle (degrees)
(angv    =                25.) Vertical viewing angle (degrees)
(command=                ) Display command
(graphic=   stdgraph) Graphics device
(default=                b) Default option for x-y input
(fixpix  =                no) Fixpix style input?
(mode    =                ql)

```

Note that the `autodisplay` option is set to `no`. This means the image only reloads to display the specified changes upon command which saves substantial time as the images are large and take tens of seconds to load. It should also be noted that it was verified that the full extent of the image was loaded and displayed for editing even though the full image size exceeds the maximum IRAF

display size; the `imedit` routine automatically applies some smoothing/binning to allow a full image to be displayed.

The major commands used for the cleaning were:

- `'b'` to replace a circular region with noise based on the surrounding background.
- `'a'` to replace a rectangular region with noise based on the surrounding background.
- `'+'` to increase the radius of the aperture used by the `'b'` command
- `'-'` to decrease the radius of the aperture used by the `'b'` command
- `'r'` to reload the image
- `':write'` to write the current image to a temporary name to protect against system crashes and mistakes in editing.

C.1.2 Shifting the Images

The default mode for interacting with IRAF is pixel coordinates so before beginning photometry it is convenient to shift the images to the same pixel units. This makes matching the photometry outputs between different images much easier. In order to align and shift the images, we rely on two routines from J. Salzer: `getshifts.cl` and `doalign.cl`. First, load these tasks into IRAF, calling them with the appropriate directory structure. Note also that the spaces around the `'='` are important.


```
ecl> task getshfts = ../getshfts.cl
ecl> task doalign = ../doalign.cl
```

After loading the tasks, setup the parameter file for `getshfts` and run the task, giving it the list of images you wish to align:

```
ecl> epar getshfts
images =          clean*
rootname=      agc198606  Text file root name
(zscale =          yes) Use zscale for display command?
(runimal=          no) Run IMALIGN is test mode?
(verbose=          yes) Verbose output?
(imageli=          )
(imexlis=          )
(mode   =          ql)
```

After running the task, the first image from the input list displays. Choose an easy to identify reference star, move the cursor to that star, and hit 'r'. Make sure to avoid choosing saturated stars as they are hard to centroid. Then choose another 10 or so reference stars by moving the cursor to them and hitting 'r'. When finished with the first image, hit 'q' to go to the next image. In subsequent images, mark only the location of the first reference star with 'r'; the location of this star in the previous image will be marked with a green box as a reference.

Next we use `doalign` to apply the shifts calculated above. This task calls `imalign`, so we first set those parameters:

```
ecl> epar imalign
input   =          Input images
referenc=          Reference image
coords  =          Reference coordinates file
output  =          Output images
(shifts =          ) Initial shifts file
```

```

(boxsize=          7) Size of the small centering box
(bigbox =         11) Size of the big centering box
(negativ=         no) Are the features negative ?
(backgro=        INDEF) Reference background level
(lower  =         INDEF) Lower threshold for data
(upper  =         INDEF) Upper threshold for data
(niterat=         3) Maximum number of iterations
(toleran=         0) Tolerance for convergence
(maxshif=        INDEF) Maximum acceptable pixel shift
(shiftim=         yes) Shift the images ?
(interp_=        poly5) Interpolant
(boundar=        constant) Boundary type
(constan=         0.) Constant for constant boundary
                        extension
(trimima=         yes) Trim the shifted images ?
(verbose=         yes) Print the centers, shifts,
                        and trim section ?

(list   =         )
(mode   =         ql)

```

After setting the parameters, quit this task and set the parameters in doalign:

```

images  =          clean*
rootname=          agc198606 Text file root name
(prefix =          sh) Prefix for shifted images
(applysh=          no) Apply the shifts?
(verbose=          yes) Verbose output?
(imageli=          )
(mode    =          ql)

```

Make sure that the `rootname` parameter matches that used in `getshifts`. Setting `prefix` to 'sh' will append a 'sh' to the front of the file names of the shifted images. The first time we set `applysh` to 'no' as we wish to check the results before shifting the images. Make sure `verbose` is set to 'yes' so that the output is printed to the screen and run the task. Check the values to make sure they are reasonable; especially check the values under 'internal' to make sure they are not

more than .05 or so. If the shifts look reasonable, run `do align`:

```
ecl> doalign app+ > shiftvalues.
```

This tells `doalign` to change the parameter `applysh` to 'yes' and to print the verbose output to a file called `shiftvalues`. You will have to hit enter a few times to verify the input parameters.

C.2 Photometric Calibration

This section describes the photometric calibration performed upon the stacked images to obtain zero points and color terms. As discussed in Chapter 5, relative photometry is used as all observed fields are located within the SDSS footprint and observations were in Sloan filters.

C.2.1 Obtaining “Standard” Stars

The first step is to obtain a list of stars from SDSS that can be used for the photometric calibration. The SkyServer Search Forum was used to locate stars within a 10' radius from the central coordinates of AGC198606 with magnitude limits of $17 < g < 20$ and $17 < i < 20$. This resulted in a list of ~60 stars. The coordinates of these stars were reformatted from decimal degrees to sexagesimal format for use with `msctvmark` in IRAF. The selected stars were marked on the i-band image and the best 20 sources (isolated, not saturated, actually stars) were selected for photometry. The selected stars were then overlaid on the g-band image to confirm that they were also good sources in that exposure; any poor detections were

removed from the photometry list.

C.2.2 Aperture Photometry of the Standard Stars

Now that we have a list of the coordinates of standard stars, we wish to perform aperture photometry to obtain the measured magnitudes of these stars. We will do the aperture photometry using the daophot package so that we are setting parameters in preparation for doing photometry of all the stars later.

```
ecl> noao
noao> digiphot
digiphot> daophot
```

First, set the parameters for all the relevant tasks:

```
PACKAGE = daophot
TASK = datapars
(scale = 1.) Image scale in units per pixel
(fwhmpsf= 6.5) FWHM of the PSF in scale units
(emissio= yes) Features are positive?
(sigma = 5.) Standard deviation of background
                in counts
(datamin= 150.) Minimum good data value
(datamax= 58000.) Maximum good data value
(noise = poisson) Noise model
(ccdread= noise) CCD readout noise image header
                keyword
(gain = gain) CCD gain image header keyword
(readnoi= 19.56) CCD readout noise in electrons
(epadu = 12.24) Gain in electrons per count
(exposur= exptime) Exposure time image header
                keyword
(airmass= airmass) Airmass image header keyword
(filter = filter) Filter image header keyword
```

```

(obstime=          ) Time of observation image header
                    keyword
(itime  =          1.) Exposure time
(xairmas=         INDEF) Airmass
(ifilter=         INDEF) Filter
(otime  =         INDEF) Time of observation
(mode   =          ql)

```

PACKAGE = daophot

TASK = centerpars

```

(calgori= centroid) Centering algorithm
(cbox   =         12.) Centering box width in scale units
(cthresh=         0.) Centering threshold in sigma above
                    background
(minsnra=         1.) Minimum signal-to-noise ratio for
                    centering algorithm
(cmaxite=        10) Maximum iterations for centering
                    algorithm
(maxshif=         1.) Maximum center shift in scale units
(clean  =         no) Symmetry clean before centering
(rclean =         1.) Cleaning radius in scale units
(rclip  =         2.) Clipping radius in scale units
(kclean =         3.) K-sigma rejection criterion in
                    skysigma
(mkcente=         no) Mark the computed center
(mode   =          ql)

```

PACKAGE = daophot

TASK = fitskypars

```

(salgori= mode) Sky fitting algorithm
(annulus=    30.) Inner radius of sky annulus in scale
                    units
(dannulu=   10.) Width of sky annulus in scale units
(skyvalu=    0.) User sky value
(smaxite=   10) Maximum number of sky fitting
                    iterations
(sloclip=    0.) Lower clipping factor in percent
(shiclip=    0.) Upper clipping factor in percent
(snrejec=   50) Maximum number of sky fitting
                    rejection iterations
(sloreje=    3.) Lower K-sigma rejection limit in sky sigma
(shireje=    3.) Upper K-sigma rejection limit in sky sigma

```

```

(khist   =    3.) Half width of histogram in sky sigma
(binsize=    0.1) Binsize of histogram in sky sigma
(smooth  =    no) Boxcar smooth the histogram
(rgrow   =    0.) Region growing radius in scale units
(mksky   =    no) Mark sky annuli on the display
(mode    =    ql)

PACKAGE = daophot
  TASK = photpars
(weighti= constant) Photometric weighting scheme
(apertur=      25) List of aperture radii in scale units
(zmag     =      0.) Zero point of magnitude scale
(mkapert=    no) Draw apertures on the display
(mode     =      ql)

```

The seeing for these images is ~ 6.5 pixels (or $0.7''$). The `scale` parameter is set to 1 so we are working in pixel units; if we wanted to work in arcseconds, `scale` would be set to 0.11 - the plate scale of pODI. Note also that the modified gain and noise terms are used so that the averaging of nine exposures is properly accounted for when calculating the error in a measured magnitude. The `min` value is set to 150; this is the minimum sky value expected for the g-band image. A higher value would be appropriate for the i-band image given the higher sky background. The `max` value is set to 58000 to well avoid any saturation issues; saturation appears to occur around 63000 counts. The `cbox` parameter is to ~ 2 times the FWHM, and an aperture of 4 times the FWHM is used. Note that the zero point of the magnitude scale is set to zero.

We are now ready to set up and run `phot`; we run it separately on the two images to avoid a mismatch in length between the number of images and the number of input coordinate files (one). Note that the input coordinate list is in *decimal degrees* and contains a list of *ONLY* R.A. and Dec (no source identification). In addition it is *VITAL* to change the `wcsin` parameter to `world`.

```

PACKAGE = daophot
TASK = phot
image = shclean_198606_g Input image(s)
coords = sdss_phot_coords.txt Input coordinate list(s)
                                (default: image.coo.?)
output = default Output photometry file(s)
                                (default: image.mag.?)
skyfile = Input sky value file(s)
(plotfil= ) Output plot metacode file
(datapar= ) Data dependent parameters
(centerp= ) Centering parameters
(fitskyp= ) Sky fitting parameters
(photpar= ) Photometry parameters
(interac= no) Interactive mode?
(radplot= no) Plot the radial profiles?
(icomman= ) Image cursor: [x y wcs]
                                key [cmd]
(gcomman= ) Graphics cursor: [x y wcs]
                                key [cmd]
(wcsin = world) The input coordinate system
                                (logical,tv,physical,world)
(wcsout = )_.wcsout) output coordinate system
                                (logical,tv,physical)
(cache = )_.cache) Cache input image pixels
                                in memory?
(verify = )_.verify) Verify critical parameters?
(update = )_.update) Update critical parameters?
(verbose= )_.verbose) Print phot messages?
(graphic= )_.graphics) Graphics device
(display= )_.display) Display device
(mode = ql)

```

C.2.3 Getting Photometric Solutions

We now wish to compare the measured magnitudes to the magnitudes from SDSS to figure out our photometric solutions. The equations we wish to solve look like:

$$g = G + ZP_G + CT_G * (G - I) \quad (C.2)$$

$$i = I + ZP_I + CT_I * (G - I), \quad (C.3)$$

where g and i are the measured magnitudes and G, I the reported SDSS magnitudes. ZP_G and ZP_I are the zeropoints of the g and i filters, respectively while CT_G and CT_I are the color terms for the two filters. Ideally, we would do this in IRAF using `fitparms`. Unfortunately, this task returns a segmentation violation. For this reason, photometric solutions were obtained in IDL using the photometric equations above after formatting the output of `phot` with `mknobsfile`:

```
PACKAGE = photcal
TASK = mknobsfile
photfile= shclean*mag* The input list of APPHOT/DAOPHOT
                        databases
idfilter= odi_g,odi_i The list of filter ids
imsets = obsfile The input image set file
observat= aper25 The output observations file
(wrap = yes) Format output for easy reading?
(obspara= ) input observing parameters file
(obscolu= 2,3,4) The format of obsparams
(minmage= 0.001) The minimum error magnitude
(shifts = ) The input x and y coordinate
                shifts file
(apercor= ) The input aperture
                corrections file
(apertur= 1) The aperture number of the
                extracted magnitude
(toleran= 10.) The tolerance in pixels for
                position matching
(allfilt= no) Output only objects matched
```



```

                                in all filters
(verbose =                       no) Verify interactive user input?
(verbose =                       yes) Print status, warning and errors?
(mode    =                       ql)

```

The final solutions are an *i* zero point magnitude of 25.83 with a color term of 0.0027 and a photometric scatter of 0.015. The *g* filter has a zero point of 26.23 with a color term of 0.11 and a photometric scatter of 0.029. Note that the color term of the *i* filter is zero within the scatter while the *g* filter has a substantial color term. This is because the *g* filter used at pODI is not an exact clone of the SDSS filters.

C.3 Photometry of the Stars

This section overviews how the stars within the image are automatically identified and have their photometry calculated. For this stage, I used the package **daophot** within IRAF and followed the directions of the manual *A Reference Guide to the IRAF/DAOPHOT Package* [Davis, 1994]. Note that it may very well be better to use **SExtractor**. One important step that is missing is to use the size of the FWHM of an object to determine if it is a star or a galaxy; **SExtractor** may provide a more natural way of doing this than **daophot**.

C.3.1 Locate the Stars

First, we use `daofind` to automatically locate all the stars in the image. The choice was made to generate a list of stars in the *i*-band image and use that list for photometry in both filters. As before, our PSF is 6.5 pixels, so set `fwhmspf` equal to

6.5 The `psfrad` is the size of the largest star we would wish to measure in the image. Visual inspection indicates this is ~ 190 pixels across, so set `psfrad` to 95. Now we will be a little more careful with our minimum count level. We calculated earlier that the sky background level was 622 with a variance of 7 counts for the i-band, so we wish to set our minimum level to be: $622 - 3 \times 7 = 600$. The parameter files should look as follows:

```

PACKAGE = daophot
  TASK = datapars
(scale = 1.) Image scale in units per pixel
(fwhmpsf= 6.5) FWHM of the PSF in scale units
(emissio= yes) Features are positive?
(sigma = 5.) Standard deviation of background
               in counts
(datamin= 600.) Minimum good data value
(datamax= 58000.) Maximum good data value
(noise = poisson) Noise model
(ccdread= noise) CCD readout noise header keyword
(gain = gain) CCD gain image header keyword
(readnoi= 19.56) CCD readout noise in electrons
(epadu = 12.24) Gain in electrons per count
(exposur= exptime) Exposure time image header keyword
(airmass= airmass) Airmass image header keyword
(filter = filter) Filter image header keyword
(obstime= ) Time of observation header keyword
(itime = 1.) Exposure time
(xairmas= INDEF) Airmass
(ifilter= INDEF) Filter
(otime = INDEF) Time of observation
(mode = ql)

PACKAGE = daophot
  TASK = daopars
(funcutio= gauss) Form of analytic component of
                  psf model
(varorde= 0) Order of empirical component of
             psf model
(nclean = 0) Number of cleaning iterations for
              computing psf model

```

```

(saturat=      no) Use wings of saturated stars in psf
                  model computation?
(matchra=      5.) Object matching radius in scale units
(psfrad =     95.) Radius of psf model in scale units
(fitrad =      7.) Fitting radius in scale units
(recente=     yes) Recenter stars during fit?
(fitsky =      no) Recompute group sky value during fit?
(groupsk=     yes) Use group rather than individual
                  sky values?
(sannulu=      0.) Inner radius of sky fitting annulus
                  in scale units
(wsannul=     11.) Width of sky fitting annulus in
                  scale units
(flaterr=     0.75) Flat field error in percent
(proferr=      5.) Profile error in percent
(maxiter=     50) Maximum number of fitting iterations
(clipexp=      6) Bad data clipping exponent
(clipran=     2.5) Bad data clipping range in sigma
(mergera=    INDEF) Critical object merging radius in
                  scale units
(critsnr=      1.) Critical S/N ratio for group
                  membership
(maxnsta=    10000) Maximum number of stars to fit
(maxgrou=      60) Maximum number of stars to fit
                  per group
(mode   =      ql)

```

```

PACKAGE = daophot
TASK = daofind

```

```

image   =   shclean_198606_i   Input image(s)
output  =      agcl98606.coo   Output coordinate file(s)
                                   (default: image.coo.?)
(starmap=                                ) Output density enhancement
                                   image(s)
(skymap =                                ) Output sky image(s)
(datapar=                                ) Data dependent parameters
(findpar=                                ) Object detection parameters
(boundar=      nearest) Boundary extension
                                   (constant|nearest|reflect|wrap)

```

```

(constan=          0.) Constant for boundary
                        extension
(interac=          no) Interactive mode?
(icomman=          ) Image cursor:
                        [x y wcs] key [cmd]
(gcomman=          ) Graphics cursor:
                        [x y wcs] key [cmd]
(wcsout  =          )_wcsout) The output coordinate system
                        (logical,tv,physical)
(cache   =          )_.cache) Cache the image pixels?
(verify  =          )_.verify) Verify critical daofind pars?
(update  =          )_.update) Update critical daofind pars?
(verbose=          )_.verbose) Print daofind messages?
(graphic=          )_.graphics) Graphics device
(display=          )_.display) Display device
(mode    =          ql)

```

```

PACKAGE = daophot
TASK    = findpars

```

```

(thresho=          5.) Threshold in sigma for
                        feature detection
(nsigma  =          1.5) Width of convolution kernel
                        in sigma
(ratio   =          1.) Ratio of minor to major
                        axis of Gaussian kernel
(theta   =          0.) Position angle of major
                        axis of Gaussian kernel
(sharplo=          0.2) Lower bound on sharpness
                        for feature detection
(sharphi=          1.) Upper bound on sharpness
                        for feature detection
(roundlo=          -1.) Lower bound on roundness
                        for feature detection
(roundhi=          1.) Upper bound on roundness
                        for feature detection
(mkdetec=          no) Mark detections on the
                        image display?
(mode    =          ql)

```

We set the variance to 7 counts and start with a threshold of 4 sigma. We mark all the found stars on the i-band image using `display` and `tvmark`. This detects a lot of stars but also a lot of junk. Most of the junk appears to be cosmic rays, meaning the manual cleaning was highly ineffective. This could be due to the large size of the images and the necessity to clean the full images, or in part to the fact that the images are somehow smoothed to allow the full image to be displayed at once during `imedit`. The threshold was gradually increased from 4 sigma to 7 sigma. This did not seem to affect the detection of stars overly much and removed some of the junk detections. I displayed the detections found at the 7 sigma threshold and manually removed any entries that appeared to be cosmic rays or have no actual counterpart in either the i-band OR g-band image. While this was much more efficient than recleaning the images to remove cosmic rays, it does introduce a strong bias to not detect faint sources near the background level. Doing a more careful job at this stage is critical in future work as the full depth of the images needs to be explored to place limits on the stellar populations out to the distances of interest. However, because one of the next steps involve PSF-fitting the detected sources it is important to not have a lot of junk detections around as the program will attempt to fit PSFs to them.

Star-Galaxy Discrimination

As emphasized in Chapter 5, the analysis of the optical images is currently strongly hampered by the poor star-galaxy separation. Visual inspection of the data reveals many detected objects that are clearly galaxies. The excellent image

quality from the WIYN telescope should allow a star-galaxy discrimination based on the FWHM of an object. It may be possible to do this within `daofind` by adjusting the sharpness and roundness parameters. Alternatively, it may be possible to calculate the FWHM of all detected objects after running `daofind` and doing the photometry below. Then, a FWHM cut could be enacted that depends on the magnitude of an object, accounting for the widening of the FWHM for low signal stars. Alternatively, **SExtractor** may provide enhanced star-galaxy separation.

C.3.2 Aperture Photometry

This step serves as a starting point for the PSF-fitting photometry of `daophot` and is also when the sky background levels are computed. We use a small aperture (\sim FWHM) for this step as we wish to ensure that we are measuring the light only from the star and not any of its neighbors. We will set some of the inputs in `phot` and make sure that `verify` is set to 'yes' so that we are prompted on the command line for the other relevant parameters.

```
PACKAGE = daophot
  TASK = phot
image    = shclean_198606_i  Input image(s)
coords   = agc198606.coo    Input coordinate list(s)
                                (default: image.coo.?)
output   =                  default Output photometry file(s)
                                (default: image.mag.?)
skyfile  =                  Input sky value file(s)
(plotfil=                  ) Output plot metacode file
(datapar=                  ) Data dependent parameters
(centerp=                  ) Centering parameters
(fitskyp=                  ) Sky fitting parameters
(photpar=                  ) Photometry parameters
(interac=                  no) Interactive mode?
```

```

(radplot=          no) Plot the radial profiles?
(icomman=          ) Image cursor: [x y wcs]
                    key [cmd]
(gcomman=          ) Graphics cursor: [x y wcs]
                    key [cmd]
(wcsin  =          )_.wcsin) The input coordinate system
                        (logical,tv,physical,world)
(wcsout =          )_.wcsout) The output coordinate system
                        (logical,tv,physical)
(cache  =          )_.cache) Cache the input image pixels?
(verify =          yes) Verify critical phot params?
(update =          yes) Update critical phot params?
(verbose=          yes) Print phot messages?
(graphic=          stdgraph) Graphics device
(display=          stdimage) Display device
(mode   =          ql)

```

Centering algorithm (none) (CR or value):

New centering algorithm: none

Sky fitting algorithm (mode) (CR or value):

Sky fitting algorithm: mode

Inner radius of sky annulus in scale units (10.)

(CR or value): 15

New inner radius of sky annulus:

15. scale units 15. pixels

Width of the sky annulus in scale units (10.)

(CR or value):

New width of the sky annulus:

10. scale units 10. pixels

File/list of aperture radii in scale units (3.)

(CR or value): 7

Aperture radius 1: 7. scale units 7. pixels

Standard deviation of background in counts (7.)

(CR or value): INDEF

New standard deviation of background: INDEF counts

Minimum good data value (600.) (CR or value):

New minimum good data value: 600. counts

Maximum good data value (60000.) (CR or value):

New maximum good data value: 60000. counts

C.3.3 Make the PSF

We need to create a model of the PSF to be used for the photometry and source subtraction. In order to do this, we select several bright, isolated stars that can be used to model the PSF. Set up and run the task `psf`:

```
PACKAGE = daophot
  TASK = psf
image    = shclean_198606_i  Input image(s) for which to
                             build PSF
photfile=          default  Input photometry file(s)
                             (default: image.mag.?)
pstfile  =          default  Input psf star list(s)
                             (default: image.pst.?)
psfimage=          default  Output PSF image(s)
                             (default: image.psf.?)
opstfile=          default  Output PSF star list(s)
                             (default: image.pst.?)
groupfil=          default  Output PSF star group file(s)
                             (default: image.psg.?)
(plotfil=          ) Output plot metacode file
(datapar=          ) Data dependent parameters
(daopars=          ) Psf fitting parameters
(matchby=          yes) Match psf star list to
                             photometry file(s) by id num?
(interac=          yes) Compute the psf interactively?
(mkstars=          no) Mark deleted and accepted psf
                             stars?
(showplo=          yes) Show plots of PSF stars?
(plottyp=          mesh) Default plot type
                             (mesh|contour|radial)
(icomman=          ) Image cursor:
                             [x y wcs] key [cmd]
(gcomman=          ) Graphics cursor:
                             [x y wcs] key [cmd]
(wcsin  =          )_.wcsin) The input coordinate system
                             (logical,tv,physical,world)
(wcsout =          )_.wcsout) The output coordinate system
                             (logical,tv,physical)
(cache  =          )_.cache) Cache the input image in memory?
```



```

(verbose =      )_.verbose) Verify critical psf parameters?
(update =      )_.update) Update critical psf parameters?
(verbose=      )_.verbose) Print psf messages?
(graphic=      stdgraph) Graphics device
(display=      stdimage) Display device
(mode   =      ql)

```

Select stars with 'a' and you will be able to view them using a variety of key commands. Hit 'a' again to accept the star as a model for constructing the PSF. When finished type 'w' to write the PSF model based on the selected stars to file and 'q' to quit.

We need to test our PSF and account for the fact that neighboring stars will impact its determination. First, we run `nstar` to fit the selected stars and their neighbors simultaneously. The parameter `psfradius` should be set to the minimum separation between one of our PSF stars and its neighbors. We will take 40 pixels to be safe. Running `nstar` on the command line we have the following interaction:

```

daophot> nstar
Image corresponding to photometry (shclean_198606_i ):
Input group file (image.grp.?) (default):
    shclean_198606_i.psg.1
PSF image (default: image.psf.?) (default):
Output photometry file (default: image.nst.?) (default):
Output rejections file (default: image.nrj.?) (default):

Recenter the stars (yes):
    Recenter the stars: yes
Refit the sky (no):
    Refit the sky: no
Use group sky values (yes):
    Use group sky values: yes
Psf radius in scale units (95.): 40
    New psf radius: 40. scale units 40. pixels

```

```

Fitting radius in scale units (7.):
    New fitting radius: 7. scale units 7. pixels
Maximum group size in number of stars (60):
    New maximum group size: 60 stars
Minimum good data value (600.) (CR or value):
    New minimum good data value: 600. counts
Maximum good data value (60000.) (CR or value):
    New maximum good data value: 60000. counts

Group:    1 contains  1 stars
  ID:  2004  XCEN:  6058.42  YCEN:  7206.28  MAG:    15.18
Group:    2 contains  2 stars
  ID:  2554  XCEN:  8339.79  YCEN:  9366.47  MAG:    16.61
  ID:  2545  XCEN:  8256.10  YCEN:  9319.55  MAG:    23.22
Group:    3 contains  3 stars
  ID:  1607  XCEN:  1573.59  YCEN:  5881.47  MAG:    15.81
  ID:  1602  XCEN:  1680.09  YCEN:  5869.78  MAG:    22.76
  ID:  1638  XCEN:  1622.13  YCEN:  5976.86  MAG:    22.67
Group:    4 contains  1 stars
  ID:  1080  XCEN:  6630.98  YCEN:  4006.65  MAG:    15.76
Group:    5 contains  1 stars
  ID:   694  XCEN:  7933.82  YCEN:  2634.98  MAG:    15.31

```

Looking at the output image, we realize that many of the psf stars subtracted out very poorly. We reselect a new set of stars with similar properties to the best subtracted star and try the steps above again. After a few iterations, we find a good set of PSF stars - the initial stars that were chosen were too big and bright, in general. Our final set of stars doesn't include any neighbors so we are done after checking our subtracted image. If our PSF stars had neighbors, we would want to create an image where the neighbors (but not PSF stars) are subtracted off, and use that image to create a final PSF.

C.3.4 Do the PSF-Fitting Photometry

We are now ready to do the PSF-fitting to the full field and obtain an initial photometry list. Set-up and run allstar:

```
PACKAGE = daophot
TASK = allstar
image = shclean_198606_i Image corresponding to
                        photometry
photfile= default Input photometry file
                        (default: image.mag.?)
psfimage= default PSF image (default:image.psf.?)
allstarf= default Output photometry file
                        (default: image.als.?)
rejfile = default Output rejections file
                        (default: image.arj.?)
subimage= default Subtracted image
                        (default: image.sub.?)
(datapar= ) Data dependent parameters
(daopars= ) Psf fitting parameters
(wcsin = )_.wcsin) The input coordinate system
                        (logical,tv,physical,world)
(wcsout = )_.wcsout) The output coordinate system
                        (logical,tv,physical)
(wcspsf = )_.wcspsf) The psf coordinate system
                        (logical,tv,physical)
(cache = yes) Cache the data in memory?
(verify = yes) Verify critical allstar pars?
(update = )_.update) Update critical allstar pars?
(verbose= )_.verbose) Print allstar messages?
(version= 2) Version
(mode = ql)
Recenter the stars (yes):
    Recenter the stars: yes
Use group sky values (yes):
    Use group sky values: yes
Refit the sky (no):
    Refit the sky: no
Psf radius in scale units (95.):
    New psf radius: 95. scale units 95. pixels
```

```
Fitting radius in scale units (7.):  
    New fitting radius: 7. scale units 7. pixels  
Maximum group size in number of stars (60):  
    New maximum group size: 60 stars  
Minimum good data value (600.) (CR or value):  
    New minimum good data value: 600. counts  
Maximum good data value (60000.) (CR or value):  
    New maximum good data value: 60000. counts
```

Examine the results in the subtracted image. We see that the PSF-fitting of the brightest stars is horrible. This could be a result of using a low value for the maximum good data value. Since these stars are foreground Milky Way stars, their poor photometry is not of concern. The fainter objects appear to be well subtracted, except for galaxies which are extended structures. A few relatively bright stars are not subtracted as these were missed by `daofind`; for this preliminary analysis we do not worry about them. There do not appear to be new sources revealed by the subtraction so we will use the list of magnitudes generated by `allstar`. If there were a number of new sources revealed by the subtraction we would wish to run `daofind`, `phot` and then `allstar` as above on the newly revealed sources, appending the photometry to the current magnitude list.

C.3.5 Aperture Correction

Since the aperture used for determining the initial photometry was small (7 pixels), we wish to calculate an aperture correction to account for the fact that we are missing light from the wings of the stars. In order to perform the aperture correction we will compare the magnitudes for a set of stars using apertures of 7 and 25 pixels. We wish to use isolated stars for this to ensure that with the larger aperture

we are only measuring light from the star of interest and not its neighbors. Since our PSF-model stars were selected for their isolation, we will use them. We use `txdump` to get the x,y coordinates of these stars for input to `phot`:

```
daophot> txdump shclean_198606_i.nst.5 xc,yc > iap
```

Now run `phot` on these stars using both apertures:

```
phot shclean_198606_i iap ires aper=7.,25. annulus=30
      dannu=10 verb+
sh*_i  6958.11  3892.12  625.5253   18.098   17.810   ok
sh*_i  1924.92  1956.34  623.8038   18.842   INDEF   err
sh*_i  2523.05  6229.58  626.3999   18.406   INDEF   err
sh*_i  7671.85  8811.21  624.8492   20.136   INDEF   err
sh*_i  7193.28  2950.06  624.8723   18.554   18.243   ok
```

For some reason we were only able to get photometry from the larger aperture for 2 of the 5 stars. We can obtain a “good enough” aperture correction using just these two stars but for future work we will want more stars with good photometry and to understand why these stars did not have magnitudes calculated for the larger aperture - is it a background issue? The average correction for the two stars is -0.3 mags.

We repeat the above steps (PSF-model creation, PSF-fitting and aperture correction) for the *g* band image. Everything is similar except that one of the PSF model stars is too faint in *g* to be used. The aperture correction values look like this:

```
txdump shclean_198606_g.nst.1 xc,yc > gap
phot shclean_198606_g gap gres aper=7.,25.
      annulus=30 dannu=10 verb+
```

sh*_g	1924.96	1956.89	240.828	20.988	20.684	ok
sh*_g	7193.11	2950.11	242.7656	20.263	20.000	ok
sh*_g	6957.75	3892.43	242.9698	19.485	19.221	ok
sh*_g	2523.26	6229.65	241.8139	18.539	18.286	ok

This time we obtain magnitudes for both apertures for all stars. We find an average aperture correction of -0.28 for the g-band. Given the similarity between the two filters and the minimal number of stars used, we choose to use -0.3 mags for both filters.

C.3.6 Transform Magnitudes to Standard System

We now wish to use our photometric coefficients from before to transform the measured magnitudes of all the stars to the standard system. Note that the zero point used to calculate the magnitudes for the whole field was reset to the default IRAF value of 25 so the zero point corrections we need to apply are small (0.83 for i and 1.23 for g). We wish to format the output from `allstar` before transforming the magnitudes. We do this with `mkobsfile` (NOT `mknobsfile`). This is also the stage at which aperture corrections are applied; setting `apercor` to `STDIN` results in a command line prompt for the aperture correction. After running `mkobsfile`, the magnitudes are transformed to the standard SDSS `ugriz` system in IDL and further analysis proceeds in IDL.

```
PACKAGE = photcal
```

```
TASK = mkobsfile
```

```
photfile=
```

```
The input list of
```

```
APPHOT/DAOPHOT databases
```

```

idfilter=      odi_g,odi_i  The list of filter ids
observat=      The output observations file
(wrap   =      yes) Format the output file for
                  easy reading ?
(imsets  =      STDIN) The input image set file
(obspara=      ) The observing parameters file
(obscolu=      2 3 4 5) The format of obsparams
(minmage=      0.001) The minimum magnitude error
(shifts  =      STDIN) The x and y coordinate shifts
(apercor=      STDIN) The aperture corrections file
(apertur=      1) The aperture number of the
                  extracted magnitude
(toleran=      5.) The tolerance in pixels for
                  position matching
(allfilt=      no) Output only objects matched
                  in all filters
(verify  =      no) Verify interactive user input ?
(verbose=      yes) Print status, warning and
                  error messages ?
(mode    =      ql)

```

C.4 Color Magnitude Diagram Analysis

The analysis of the detected sources is done by applying a color magnitude diagram (CMD) filter cut to all sources based on an old, metal poor stellar pop-

ulation for a given distance modulus. The remaining sources are then spatially smoothed to enhance any overdensity. Section 5.5 describes the motivation behind the methodology in detail. The following IDL code contains the full details on how this procedure is done:

pro cm

```
;This procedure applies color magnitude filters
;to the observed sources and then does
;spatial smoothing

;First read in the isochrones, which live in
;/home/maricao/betsey/uchvcs
restore, '/home/maricao/betsey/uchvcs/isochrone_template.sav'
dir='/home/maricao/betsey/uchvcs/'
file_0006_8gyr = 'isochrone_ez0006_8gyr_av085.dat'
file_0001_8gyr = 'isochrone_z0001_8gyr_av085.dat'
file_0006_135gyr = 'isochrone_z0006_13.5gyr_av085.dat'
file_0001_135gyr = 'isochrone_z0001_13.5gyr_av085.dat'
iso_0006_8gyr=read_ascii(dir+file_0006_8gyr,$
                        template=isochrone_template)
iso_0001_8gyr=read_ascii(dir+file_0001_8gyr,$
                        template=isochrone_template)
iso_0006_135gyr=read_ascii(dir+file_0006_135gyr,$
                        template=isochrone_template)
iso_0001_135gyr=read_ascii(dir+file_0001_135gyr,$
                        template=isochrone_template)

;Get the colors of the isochrones
gicolor_0006_8gyr = iso_0006_8gyr.g - iso_0006_8gyr.i
gicolor_0001_8gyr = iso_0001_8gyr.g - iso_0001_8gyr.i
gicolor_0006_135gyr = iso_0006_135gyr.g - iso_0006_135gyr.i
gicolor_0001_135gyr = iso_0001_135gyr.g - iso_0001_135gyr.i

;Find the stars that are appropriate to use for the filter.
;Specifically, we don't want stars that have evolved
;off the RGB to white dwarfs, etc.
;Evolution off the RGB/AGB is represented by excessive
;mass loss compared to the initial mass.
;Empirically determined the following limits by
;plotting CMDs for the isochrones and placing limits
;until evolution to WD is not seen.
```



```

good_0006_8gyr = where( (isochrone_0006_8gyr.m_act / $
                        isochrone_0006_8gyr.m_ini) ge 0.65)
good_0001_8gyr = where( (isochrone_0001_8gyr.m_act / $
                        isochrone_0001_8gyr.m_ini) ge 0.7)
good_0006_135gyr=where((isochrone_0006_135gyr.m_act / $
                        isochrone_0006_135gyr.m_ini) ge 0.7)
good_0001_135gyr=where((isochrone_0001_135gyr.m_act / $
                        isochrone_0001_135gyr.m_ini) ge 0.75)
;Create master arrays storing the "good"
;isochrone values that will be used
master_i = [isochrone_0006_8gyr.i[good_0006_8gyr],$
            isochrone_0001_8gyr.i[good_0001_8gyr],$
            isochrone_0006_135gyr.i[good_0006_135gyr],$
            isochrone_0001_135gyr.i[good_0001_135gyr]]
master_gi = [gicolor_0006_8gyr[good_0006_8gyr],$
            gicolor_0001_8gyr[good_0001_8gyr],$
            gicolor_0006_135gyr[good_0006_135gyr],$
            gicolor_0001_135gyr[good_0001_135gyr]]

;Set up a grid that ranges in color from
;(-1,3) and mag (24,17)
;increment by 0.05
;dimensions are (4*0.05+1) and (7*.05+1)
param_space = fltarr(4/0.05+1,7/.05+1)
x_index = findgen(4/.05+1)*.05 -1.
y_index = -findgen(7/.05+1)*.05+24.
;and record the size of the grid
nx=n_elements(x_index)
ny=n_elements(y_index)

;now restore the arrays containing information on
;the detected objects.
restore,'stars_198606.sav'
;And give the i and g arrays better names so that
;loops can be indexed using i.
i_obs = i
g_obs = g
;Set limits on what objects are considered well measured
;Define the color and an error based on the reported errors
;in individual filters
color=g-i
cerr=sqrt(ierr^2 + gerr^2)

```

```

;Only keep objects with errors less than 0.2 mags
good_stars = where(cerr le 0.2)
;And use the same color error for widening the filter
sigma_gi = 0.2

;Now do all the steps (filtering , smoothing , etc..)
;for distance modulus =22.4 as an example
;Code can be adapted to any distance modulus
;in a straight-forward manner

;First set the master array to have
;the magnitudes appropriate for the distance modulus
master_i_224 = master_i + 22.4
;and create a parameter space array specifically for this
;distance modulus
param_space_224 = param_space

;Now create the filter by looping through the
;parameter space array and checking if that
;magnitude/color combination falls within
;the color error point of an isochrone.
for i=0,nx-1 do begin
    for j=0,ny-1 do begin
        ;get mag and color at this location
        imag = y_index[j]
        color = x_index[i]
;check if this combination is contained within an
;isochrone value
        check = where( abs(master_gi-color) le sigma_gi AND $
                        abs(imag - master_i_224) le sigma_gi,$
                        cnt_isochrone)
;if it does , update param_space to reflect that
        if cnt_isochrone ge 1 then param_space_224[i,j] = 1.
    endfor
endfor

;Now use the param_space_224 as a filter .
;Iterate through each star/source -
;find the index of its color and i mag relative to
;the param_space array.
;Then check if the array lets it pass or not ,
;and if it does , flag it as passed

```

```

ns=n_elements(i_obs)
check_star_224 = fltarr(ns)
gi_obs = g_obs-i_obs
for j=0,ns-1 do begin
    mincolor= min( abs(x_index - gi_obs[j]),ind_param_color)
    minmag = min( abs(y_index - i_obs[j]),ind_param_mag)
    check_cut=param_space_224[ind_param_color,ind_param_mag]
    ;if star passes cut pass it a save flag
    if check_cut eq 1 then check_star_224[j] = 1
endfor

;now all the stars that pass cut should have check_star = 1
pass = where(check_star_224 eq 1)

;Now it is time to smooth the stars
;First, the stars need to be gridded
;Degrade the resolution by a factor of 10
;for computational reasons
;This is a grid separation of 1.1" which
;should be plenty
grid_10res = fltarr(1101,1101)
;stay in pixel values
x_array_pix = findgen(1101)
y_array_pix = findgen(1101)

;Go through all the stars and add 1 to each point in
;grid_10res that contains a star.
n=n_elements(pass)
for j=0,n-1 do begin
    ;find the location of the star in degraded coordinates
    xtmp = round(x[pass[j]]/10.)
    ytmp = round(y[pass[j]]/10.)
    ;increment the grid at that location
    grid_10res[xtmp,ytmp] = grid_10res[xtmp,ytmp]+1.
endfor

;now smooth to all the resolutions of interest
n=n_elements(grid_10res)
;first create the kernel
print, 'Creating a 4arcmin kernel'
gausskernel4=psf_gaussian(npixel=1101,fwhm=218,ndimen=2,$

```

```

                                /normalize , centroid =[0 ,0])
;then smooth via the FFT and convolution theorem
print , 'Smoothing the grid to 4'
grid_smooth4=real_part(FFT( fft ( grid_10res )*$
                                fft ( gausskernel4 ),/ inverse ))/$
                                sqrt(n)
;Repeat for the other 3 smoothing scales
print , 'Creating a 3arcmin kernel'
gausskernel3=psf_gaussian ( npixel=1101 , fwhm=164 , ndimen=2 , $
                                /normalize , centroid =[0 ,0])
print , 'Smoothing the grid to 3'
grid_smooth3=real_part(FFT( fft ( grid_10res )*$
                                fft ( gausskernel3 ),/ inverse ))/$
                                sqrt(n)
print , 'Creating a 2arcmin kernel'
gausskernel2=psf_gaussian ( npixel=1101 , fwhm=109 , ndimen=2 , $
                                /normalize , centroid =[0 ,0])
print , 'Smoothing the grid to 2'
grid_smooth2=real_part(FFT( fft ( grid_10res )*$
                                fft ( gausskernel2 ),/ inverse ))/$
                                sqrt(n)
print , 'Creating a 1arcmin kernel'
gausskernel1=psf_gaussian ( npixel=1101 , fwhm=55 , ndimen=2 , $
                                /normalize , centroid =[0 ,0])
print , 'Smoothing the grid to 1'
grid_smooth1=real_part(FFT( fft ( grid_10res )*$
                                fft ( gausskernel1 ),/ inverse ))/$
                                sqrt(n)

;plot results as desired and this is the end

end

```

BIBLIOGRAPHY

- R. Barkana and A. Loeb. The physics and early history of the intergalactic medium. *Reports on Progress in Physics*, 70:627–657, April 2007. doi: 10.1088/0034-4885/70/4/R02.
- D. G. Barnes and W. J. G. de Blok. The neutral hydrogen environments of the nearby galaxies WLM, NGC 1313 and Sextans A. *MNRAS*, 351:333–338, June 2004. doi: 10.1111/j.1365-2966.2004.07790.x.
- A. Begum and J. N. Chengalur. Kinematics of the faintest gas-rich galaxy in the Local Group: DDO210. *A&A*, 413:525–534, January 2004. doi: 10.1051/0004-6361:20031549.
- A. Begum, J. N. Chengalur, I. D. Karachentsev, and M. E. Sharina. Resolving the mystery of the dwarf galaxy HIZSS003. *MNRAS*, 359:L53–L57, May 2005. doi: 10.1111/j.1745-3933.2005.00040.x.
- A. Begum, J. N. Chengalur, I. D. Karachentsev, S. S. Kaisin, and M. E. Sharina. Gas distribution, kinematics and star formation in faint dwarf galaxies. *MNRAS*, 365:1220–1234, February 2006. doi: 10.1111/j.1365-2966.2005.09817.x.
- A. Begum, S. Stanimirović, J. E. Peek, N. P. Ballering, C. Heiles, K. A. Douglas, M. Putman, S. J. Gibson, J. Grcevich, E. J. Korpela, M.-Y. Lee, D. Saul, and J. S. Gallagher, III. Compact H I Clouds from the GALFA-H I Survey. *ApJ*, 722:395, October 2010. doi: 10.1088/0004-637X/722/1/395.
- E. Z. Bernstein-Cooper, J. M. Cannon, R. Giovanelli, M. P. Haynes, E. A. K. Adams, K. L. Rhode, J. J. Salzer, E. D. Skillman, , and K. B. W. McQuinn. High Resolution HI Observations of Leo P. in preparation, 2013.
- G. Besla, N. Kallivayalil, L. Hernquist, R. P. van der Marel, T. J. Cox, and D. Kereš. Simulations of the Magellanic Stream in a First Infall Scenario. *ApJL*, 721:L97–L101, October 2010. doi: 10.1088/2041-8205/721/2/L97.
- B. Binggeli, C. C. Popescu, and G. A. Tammann. The kinematics of the Virgo cluster revisited. *A&AS*, 98:275–296, April 1993.
- S. Blais-Ouellette, P. Amram, and C. Carignan. Accurate Determination of the Mass Distribution in Spiral Galaxies. II. Testing the Shape of Dark Halos. *AJ*, 121:1952–1964, April 2001. doi: 10.1086/319944.
- M. R. Blanton, R. H. Lupton, D. J. Schlegel, M. A. Strauss, J. Brinkmann, M. Fukugita, and J. Loveday. The Properties and Luminosity Function of Extremely Low Luminosity Galaxies. *ApJ*, 631:208–230, September 2005. doi: 10.1086/431416.

- L. Blitz, D. N. Spergel, P. J. Teuben, D. Hartmann, and W. B. Burton. High-Velocity Clouds: Building Blocks of the Local Group. *ApJ*, 514:818–843, April 1999. doi: 10.1086/306963.
- M. S. Bovill and M. Ricotti. Where are the Fossils of the First Galaxies? I. Local Volume Maps and Properties of the Undetected Dwarfs. *ApJ*, 741:17, November 2011. doi: 10.1088/0004-637X/741/1/17.
- J. Bovy and C. Dvorkin. Low-mass Suppression of the Satellite Luminosity Function Due to the Supersonic Baryon-Cold-dark-matter Relative Velocity. *ApJ*, 768:70, May 2013. doi: 10.1088/0004-637X/768/1/70.
- M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat. Too big to fail? The puzzling darkness of massive Milky Way subhaloes. *MNRAS*, 415:L40–L44, July 2011. doi: 10.1111/j.1745-3933.2011.01074.x.
- M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat. The Milky Way’s bright satellites as an apparent failure of Λ CDM. *MNRAS*, 422:1203–1218, May 2012. doi: 10.1111/j.1365-2966.2012.20695.x.
- R. Braun and W. B. Burton. The kinematic and spatial deployment of compact, isolated high-velocity clouds. *A&A*, 341:437–450, January 1999.
- R. Braun and W. B. Burton. Status of H I searches for CHVCs beyond the Local Group. *A&A*, 375:219–226, August 2001. doi: 10.1051/0004-6361:20010810.
- D. Briggs. *High Fidelity Deconvolution of Moderately Resolved Sources*. PhD thesis.
- A. M. Brooks, M. Kuhlen, A. Zolotov, and D. Hooper. A Baryonic Solution to the Missing Satellites Problem. *ApJ*, 765:22, March 2013. doi: 10.1088/0004-637X/765/1/22.
- T. M. Brown, J. Tumlinson, M. Geha, E. N. Kirby, D. A. VandenBerg, R. R. Muñoz, J. S. Kalirai, J. D. Simon, R. J. Avila, P. Guhathakurta, A. Renzini, and H. C. Ferguson. The Primeval Populations of the Ultra-faint Dwarf Galaxies. *ApJL*, 753:L21, July 2012. doi: 10.1088/2041-8205/753/1/L21.
- C. Brüns and T. Westmeier. H I observations of an Ultra-Compact High-Velocity Cloud. *A&A*, 426:L9–L13, October 2004. doi: 10.1051/0004-6361:200400068.
- J. M. Cannon, R. Giovanelli, M. P. Haynes, S. Janowiecki, A. Parker, J. J. Salzer, E. A. K. Adams, E. Engstrom, S. Huang, K. B. W. McQuinn, J. Ott, A. Saintonge, E. D. Skillman, J. Allan, G. Erny, P. Fliss, and A. Smith. The Survey of H I in Extremely Low-mass Dwarfs (SHIELD). *ApJL*, 739:L22, September 2011. doi: 10.1088/2041-8205/739/1/L22.

- K. M. Chynoweth, G. I. Langston, K. Holley-Bockelmann, and F. J. Lockman. A Search for Diffuse Neutral Hydrogen and H I Clouds in the NGC 2403 Group. *AJ*, 138:287–294, July 2009. doi: 10.1088/0004-6256/138/1/287.
- K. M. Chynoweth, K. Holley-Bockelmann, E. Polisensky, and G. I. Langston. The Origin of Neutral Hydrogen Clouds in Nearby Galaxy Groups: Exploring the Range of Galaxy Interactions. *AJ*, 142:137, October 2011a. doi: 10.1088/0004-6256/142/4/137.
- K. M. Chynoweth, G. I. Langston, and K. Holley-Bockelmann. H I Clouds in the M81 Filament as Dark Matter Minihalos – a Phase-space Mismatch. *AJ*, 141:9, January 2011b. doi: 10.1088/0004-6256/141/1/9.
- T. W. Connors, D. Kawata, and B. K. Gibson. N-body simulations of the Magellanic stream. *MNRAS*, 371:108–120, September 2006. doi: 10.1111/j.1365-2966.2006.10659.x.
- K. H. Cook, M. Mateo, E. W. Olszewski, S. S. Vogt, C. Stubbs, and A. Diercks. The Systemic Velocity and Internal Kinematics of the Dwarf Galaxy LGS 3: an Optical Foray beyond the Milky Way. *PASP*, 111:306–312, March 1999. doi: 10.1086/316339.
- T. J. Cornwell. Multiscale CLEAN Deconvolution of Radio Synthesis Images. *IEEE Journal of Selected Topics in Signal Processing*, 2:793–801, November 2008. doi: 10.1109/JSTSP.2008.2006388.
- S. Courteau and S. van den Bergh. The Solar Motion Relative to the Local Group. *AJ*, 118:337–345, July 1999. doi: 10.1086/300942.
- L. E. Davis. *A Reference Guide to the IRAF/DAOPHOT Package*. IRAF Programming Group, January 1994.
- W. J. G. de Blok and A. Bosma. High-resolution rotation curves of low surface brightness galaxies. *A&A*, 385:816–846, April 2002. doi: 10.1051/0004-6361:20020080.
- W. J. G. de Blok, S. S. McGaugh, A. Bosma, and V. C. Rubin. Mass Density Profiles of Low Surface Brightness Galaxies. *ApJL*, 552:L23–L26, May 2001a. doi: 10.1086/320262.
- W. J. G. de Blok, S. S. McGaugh, and V. C. Rubin. High-Resolution Rotation Curves of Low Surface Brightness Galaxies. II. Mass Models. *AJ*, 122:2396–2427, November 2001b. doi: 10.1086/323450.

- W. J. G. de Blok, M. A. Zwaan, M. Dijkstra, F. H. Briggs, and K. C. Freeman. An H I survey of the Centaurus and Sculptor groups. Constraints on the space density of low mass galaxies. *A&A*, 382:43–52, January 2002. doi: 10.1051/0004-6361:20011553.
- W. J. G. de Blok, A. Bosma, and S. McGaugh. Simulating observations of dark matter dominated galaxies: towards the optimal halo profile. *MNRAS*, 340:657–678, April 2003. doi: 10.1046/j.1365-8711.2003.06330.x.
- V. de Heij, R. Braun, and W. B. Burton. An automated search for compact high-velocity clouds in the Leiden/Dwingeloo Survey. *A&A*, 391:159–178, August 2002a. doi: 10.1051/0004-6361:20020870.
- V. de Heij, R. Braun, and W. B. Burton. An all-sky study of compact, isolated high-velocity clouds. *A&A*, 392:417–451, September 2002b. doi: 10.1051/0004-6361:20020908.
- S. De Rijcke, P. Prugniel, F. Simien, and H. Dejonghe. The internal dynamics of the Local Group dwarf elliptical galaxies NGC 147, 185 and 205. *MNRAS*, 369: 1321–1333, July 2006. doi: 10.1111/j.1365-2966.2006.10377.x.
- L. Dedes and P. W. M. Kalberla. Properties of extra-planar H I clouds in the outer part of the Milky Way. *A&A*, 509:A60, January 2010. doi: 10.1051/0004-6361/200912673.
- J. Diemand, M. Kuhlen, P. Madau, M. Zemp, B. Moore, D. Potter, and J. Stadel. Clumps and streams in the local dark matter distribution. *Nature*, 454:735–738, August 2008. doi: 10.1038/nature07153.
- N. H. Dieter. Neutral hydrogen near the galactic poles. *AJ*, 70:552, October 1965. doi: 10.1086/109780.
- Y. Faerman, A. Sternberg, and C. F. McKee. Ultra-Compact High Velocity Clouds as Minihalos. submitted, 2013.
- A. Ferrara and E. Tolstoy. The role of stellar feedback and dark matter in the evolution of dwarf galaxies. *MNRAS*, 313:291–309, April 2000. doi: 10.1046/j.1365-8711.2000.03209.x.
- A. Fialkov, R. Barkana, D. Tseliakhovich, and C. M. Hirata. Impact of the relative motion between the dark matter and baryons on the first stars: semi-analytical modelling. *MNRAS*, 424:1335–1345, August 2012. doi: 10.1111/j.1365-2966.2012.21318.x.
- H. A. Ford, F. J. Lockman, and N. M. McClure-Griffiths. Milky Way Disk-Halo Transition in H I: Properties of the Cloud Population. *ApJ*, 722:367–379, October 2010. doi: 10.1088/0004-637X/722/1/367.

- M. Fukugita and P. J. E. Peebles. Massive Coronae of Galaxies. *ApJ*, 639:590–599, March 2006. doi: 10.1086/499556.
- M. Geha, M. R. Blanton, M. Masjedi, and A. A. West. The Baryon Content of Extremely Low Mass Dwarf Galaxies. *ApJ*, 653:240–254, December 2006. doi: 10.1086/508604.
- M. Geha, R. P. van der Marel, P. Guhathakurta, K. M. Gilbert, J. Kalirai, and E. N. Kirby. Local Group Dwarf Elliptical Galaxies. II. Stellar Kinematics to Large Radii in NGC 147 and NGC 185. *ApJ*, 711:361–373, March 2010. doi: 10.1088/0004-637X/711/1/361.
- R. Giovanelli, M. P. Haynes, B. R. Kent, P. Perillat, A. Saintonge, N. Brosch, B. Catinella, G. L. Hoffman, S. Stierwalt, K. Spekkens, M. S. Lerner, K. L. Masters, E. Momjian, J. L. Rosenberg, C. M. Springob, A. Boselli, V. Charmandaris, J. K. Darling, J. Davies, D. Garcia Lambas, G. Gavazzi, C. Giovanardi, E. Hardy, L. K. Hunt, A. Iovino, I. D. Karachentsev, V. E. Karachentseva, R. A. Koopmann, C. Marinoni, R. Minchin, E. Muller, M. Putman, C. Pantoja, J. J. Salzer, M. Scodeggio, E. Skillman, J. M. Solanes, C. Valotto, W. van Driel, and L. van Zee. The Arecibo Legacy Fast ALFA Survey. I. Science Goals, Survey Design, and Strategy. *AJ*, 130:2598–2612, December 2005. doi: 10.1086/497431.
- R. Giovanelli, M. P. Haynes, B. R. Kent, A. Saintonge, S. Stierwalt, A. Altaf, T. Balonek, N. Brosch, S. Brown, B. Catinella, A. Furniss, J. Goldstein, G. L. Hoffman, R. A. Koopmann, D. A. Kornreich, B. Mahmood, A. M. Martin, K. L. Masters, A. Mitschang, E. Momjian, P. H. Nair, J. L. Rosenberg, and B. Walsh. The Arecibo Legacy Fast ALFA Survey. III. H I Source Catalog of the Northern Virgo Cluster Region. *AJ*, 133:2569–2583, June 2007. doi: 10.1086/516635.
- R. Giovanelli, M. P. Haynes, B. R. Kent, and E. A. K. Adams. Are Newly Discovered H I High-Velocity Clouds Minihalos in the Local Group? *ApJL*, 708: L22–L25, January 2010. doi: 10.1088/2041-8205/708/1/L22.
- R. Giovanelli, M. P. Haynes, E. A. K. Adams, J. M. Cannon, K. L. Rhode, J. J. Salzer, E. D. Skillman, E. Z. Bernstein-Cooper, and K. B. W. McQuinn. ALFALFA Discovery of the Nearby Gas-rich Dwarf Galaxy Leo P. I. H I Observations. *AJ*, 146:15, July 2013. doi: 10.1088/0004-6256/146/1/15.
- L. Girardi, E. K. Grebel, M. Odenkirchen, and C. Chiosi. Theoretical isochrones in several photometric systems. II. The Sloan Digital Sky Survey ugriz system. *A&A*, 422:205–215, July 2004. doi: 10.1051/0004-6361:20040250.
- O. Gnat, A. Sternberg, Y. Faerman, C. F. McKee, E. O. Ofek, E. A. K. Adams, R. Giovanelli, and M. Haynes. Are the Ultra-Compact High-Velocity Clouds Minihalos? Constraints from Quasar Absorption Lines. in preparation, 2013.

- F. Governato, A. Zolotov, A. Pontzen, C. Christensen, S. H. Oh, A. M. Brooks, T. Quinn, S. Shen, and J. Wadsley. Cuspy no more: how outflows affect the central dark matter and baryon distribution in Λ cold dark matter galaxies. *MNRAS*, 422:1231–1240, May 2012. doi: 10.1111/j.1365-2966.2012.20696.x.
- J. Grcevich and M. E. Putman. H I in Local Group Dwarf Galaxies and Stripping by the Galactic Halo. *ApJ*, 696:385–395, May 2009. doi: 10.1088/0004-637X/696/1/385.
- E. K. Grebel and J. S. Gallagher, III. The Impact of Reionization on the Stellar Populations of Nearby Dwarf Galaxies. *ApJL*, 610:L89–L92, August 2004. doi: 10.1086/423339.
- D. Hartmann and W. B. Burton. *Atlas of Galactic Neutral Hydrogen*. February 1997.
- M. P. Haynes, R. Giovanelli, A. M. Martin, K. M. Hess, A. Saintonge, E. A. K. Adams, G. Hallenbeck, G. L. Hoffman, S. Huang, B. R. Kent, R. A. Koopmann, E. Papastergis, S. Stierwalt, T. J. Balonek, D. W. Craig, S. J. U. Higdon, D. A. Kornreich, J. R. Miller, A. A. O’Donoghue, R. P. Olowin, J. L. Rosenberg, K. Spekkens, P. Troischt, and E. M. Wilcots. The Arecibo Legacy Fast ALFA Survey: The α .40 H I Source Catalog, Its Characteristics and Their Impact on the Derivation of the H I Mass Function. *AJ*, 142:170, November 2011. doi: 10.1088/0004-6256/142/5/170.
- M. Hoeft and S. Gottlöber. Dwarf Galaxies in Voids: Dark Matter Halos and Gas Cooling. *Advances in Astronomy*, 2010:693968, 2010. doi: 10.1155/2010/693968.
- M. Hoeft, G. Yepes, S. Gottlöber, and V. Springel. Dwarf galaxies in voids: suppressing star formation with photoheating. *MNRAS*, 371:401–414, September 2006. doi: 10.1111/j.1365-2966.2006.10678.x.
- G. L. Hoffman, E. E. Salpeter, B. Farhat, T. Roos, H. Williams, and G. Helou. Arecibo H I Mapping of a Large Sample of Dwarf Irregular Galaxies. *ApJS*, 105:269, August 1996. doi: 10.1086/192314.
- R. Ibata, N. F. Martin, M. Irwin, S. Chapman, A. M. N. Ferguson, G. F. Lewis, and A. W. McConnachie. The Haunted Halos of Andromeda and Triangulum: A Panorama of Galaxy Formation in Action. *ApJ*, 671:1591–1623, December 2007. doi: 10.1086/522574.
- J. A. Irwin, G. L. Hoffman, K. Spekkens, M. P. Haynes, R. Giovanelli, S. M. Linder, B. Catinella, E. Momjian, B. S. Koribalski, J. Davies, E. Brinks, W. J. G. de Blok, M. E. Putman, and W. van Driel. Λ CDM Satellites and H I Companions – the Arecibo ALFA Survey of NGC 2903. *ApJ*, 692:1447–1463, February 2009. doi: 10.1088/0004-637X/692/2/1447.

- I. D. Karachentsev and D. A. Makarov. The Galaxy Motion Relative to Nearby Galaxies and the Local Velocity Field. *AJ*, 111:794, February 1996. doi: 10.1086/117825.
- G. Kauffmann, S. D. M. White, and B. Guiderdoni. The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. *MNRAS*, 264:201, September 1993.
- B. R. Kent, R. Giovanelli, M. P. Haynes, A. M. Martin, A. Saintonge, S. Stierwalt, T. J. Balonek, N. Brosch, and R. A. Koopmann. The Arecibo Legacy Fast Alfa Survey. VI. Second HI Source Catalog of the Virgo Cluster Region. *AJ*, 136:713–724, August 2008. doi: 10.1088/0004-6256/136/2/713.
- A. A. Kepley, E. M. Wilcots, D. A. Hunter, and T. Nordgren. A High-Resolution Study of the H I Content of Local Group Dwarf Irregular Galaxy WLM. *AJ*, 133: 2242–2257, May 2007. doi: 10.1086/513716.
- E. N. Kirby, J. G. Cohen, and M. Bellazzini. The Dynamics and Metallicity Distribution of the Distant Dwarf Galaxy VV124. *ApJ*, 751:46, May 2012. doi: 10.1088/0004-637X/751/1/46.
- A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada. Where Are the Missing Galactic Satellites? *ApJ*, 522:82–92, September 1999. doi: 10.1086/307643.
- K. Kovač, T. A. Oosterloo, and J. M. van der Hulst. A blind HI survey in the Canes Venatici region. *MNRAS*, 400:743–765, December 2009. doi: 10.1111/j.1365-2966.2009.14662.x.
- A. Kravtsov. Dark Matter Substructure and Dwarf Galactic Satellites. *Advances in Astronomy*, 2010:281913, 2010. doi: 10.1155/2010/281913.
- F. J. Lockman. Discovery of a Population of H I Clouds in the Galactic Halo. *ApJL*, 580:L47–L50, November 2002. doi: 10.1086/345495.
- F. J. Lockman and Y. Pidopryhora. Extra-planar H I in the Inner Milky Way. In R. Braun, editor, *Extra-Planar Gas*, volume 331 of *Astronomical Society of the Pacific Conference Series*, page 59, June 2005.
- E. Ł. Łokas. The mass and velocity anisotropy of the Carina, Fornax, Sculptor and Sextans dwarf spheroidal galaxies. *MNRAS*, 394:L102–L106, March 2009. doi: 10.1111/j.1745-3933.2009.00620.x.
- U. Maio, L. V. E. Koopmans, and B. Ciardi. The impact of primordial supersonic flows on early structure formation, reionization and the lowest-mass dwarf galaxies. *MNRAS*, 412:L40–L44, March 2011. doi: 10.1111/j.1745-3933.2010.01001.x.

- P. R. Maloney and M. E. Putman. Are Compact High-Velocity Clouds Extragalactic Objects? *ApJ*, 589:270–280, May 2003. doi: 10.1086/374547.
- D. Marchesini, E. D’Onghia, G. Chincarini, C. Firmani, P. Conconi, E. Molinari, and A. Zacchei. $H\alpha$ Rotation Curves: The Soft Core Question. *ApJ*, 575:801–813, August 2002. doi: 10.1086/341475.
- A. M. Martin, E. Papastergis, R. Giovanelli, M. P. Haynes, C. M. Springob, and S. Stierwalt. The Arecibo Legacy Fast ALFA Survey. X. The H I Mass Function and Ω_{HI} from the 40% ALFALFA Survey. *ApJ*, 723:1359–1374, November 2010. doi: 10.1088/0004-637X/723/2/1359.
- N. F. Martin, J. T. A. de Jong, and H.-W. Rix. A Comprehensive Maximum Likelihood Analysis of the Structural Properties of Faint Milky Way Satellites. *ApJ*, 684:1075–1092, September 2008. doi: 10.1086/590336.
- S. Mashchenko, J. Wadsley, and H. M. P. Couchman. Stellar Feedback in Dwarf Galaxy Formation. *Science*, 319:174–, January 2008. doi: 10.1126/science.1148666.
- M. L. Mateo. Dwarf Galaxies of the Local Group. *ARA&A*, 36:435–506, 1998. doi: 10.1146/annurev.astro.36.1.435.
- D. S. Mathewson, M. N. Cleary, and J. D. Murray. The Magellanic stream. *ApJ*, 190:291–296, June 1974. doi: 10.1086/152875.
- A. W. McConnachie. The Observed Properties of Dwarf Galaxies in and around the Local Group. *AJ*, 144:4, July 2012. doi: 10.1088/0004-6256/144/1/4.
- A. W. McConnachie, M. J. Irwin, R. A. Ibata, J. Dubinski, L. M. Widrow, N. F. Martin, P. Côté, A. L. Dotter, J. F. Navarro, A. M. N. Ferguson, T. H. Puzia, G. F. Lewis, A. Babul, P. Barmby, O. Bienaymé, S. C. Chapman, R. Cockcroft, M. L. M. Collins, M. A. Fardal, W. E. Harris, A. Huxor, A. D. Mackey, J. Peñarrubia, R. M. Rich, H. B. Richer, A. Siebert, N. Tanvir, D. Valls-Gabaud, and K. A. Venn. The remnants of galaxy formation from a panoramic survey of the region around M31. *Nature*, 461:66–69, September 2009. doi: 10.1038/nature08327.
- S. S. McGaugh. The Baryonic Tully-Fisher Relation of Gas-rich Galaxies as a Test of Λ CDM and MOND. *AJ*, 143:40, February 2012. doi: 10.1088/0004-6256/143/2/40.
- S. S. McGaugh, V. C. Rubin, and W. J. G. de Blok. High-Resolution Rotation Curves of Low Surface Brightness Galaxies. I. Data. *AJ*, 122:2381–2395, November 2001. doi: 10.1086/323448.

- M. J. Meyer, M. A. Zwaan, R. L. Webster, L. Staveley-Smith, E. Ryan-Weber, M. J. Drinkwater, D. G. Barnes, M. Howlett, V. A. Kilborn, J. Stevens, M. Waugh, M. J. Pierce, R. Bhathal, W. J. G. de Blok, M. J. Disney, R. D. Ekers, K. C. Freeman, D. A. Garcia, B. K. Gibson, J. Harnett, P. A. Henning, H. Jerjen, M. J. Kesteven, P. M. Knezek, B. S. Koribalski, S. Mader, M. Marquarding, R. F. Minchin, J. O'Brien, T. Oosterloo, R. M. Price, M. E. Putman, S. D. Ryder, E. M. Sadler, I. M. Stewart, F. Stootman, and A. E. Wright. The HIPASS catalogue - I. Data presentation. *MNRAS*, 350:1195–1209, June 2004. doi: 10.1111/j.1365-2966.2004.07710.x.
- J. C. Mihos, K. M. Keating, K. Holley-Bockelmann, D. J. Pisano, and N. E. Kassim. The H I Environment of the M101 Group. *ApJ*, 761:186, December 2012. doi: 10.1088/0004-637X/761/2/186.
- R. F. Minchin, M. J. Disney, P. J. Boyce, W. J. G. de Blok, Q. A. Parker, G. D. Banks, K. C. Freeman, D. A. Garcia, B. K. Gibson, M. Grossi, R. F. Haynes, P. M. Knezek, R. H. Lang, D. F. Malin, R. M. Price, I. M. Stewart, and A. E. Wright. HIDEEP - an extragalactic blind survey for very low column-density neutral hydrogen. *MNRAS*, 346:787–802, December 2003. doi: 10.1111/j.1365-2966.2003.07134.x.
- B. Moore. Evidence against dissipation-less dark matter from observations of galaxy haloes. *Nature*, 370:629–631, August 1994. doi: 10.1038/370629a0.
- B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel, and P. Tozzi. Dark Matter Substructure within Galactic Halos. *ApJL*, 524:L19–L22, October 1999. doi: 10.1086/312287.
- R. R. Muñoz, M. Geha, and B. Willman. Turning the Tides on the Ultra-faint Dwarf Spheroidal Galaxies: Coma Berenices and Ursa Major II. *AJ*, 140:138–151, July 2010. doi: 10.1088/0004-6256/140/1/138.
- C. A. Muller, J. H. Oort, and E. Raimond. Hydrogène neutre dans la couronne galactique? *Academie des Sciences Paris Comptes Rendus*, 257:1661–1662, 1963.
- J. F. Navarro, V. R. Eke, and C. S. Frenk. The cores of dwarf galaxy haloes. *MNRAS*, 283:L72–L78, December 1996.
- D. L. Nidever, S. R. Majewski, and W. B. Burton. The Origin of the Magellanic Stream and Its Leading Arm. *ApJ*, 679:432–459, May 2008. doi: 10.1086/587042.
- D. L. Nidever, S. R. Majewski, W. Butler Burton, and L. Nigra. The 200° Long Magellanic Stream System. *ApJ*, 723:1618–1631, November 2010. doi: 10.1088/0004-637X/723/2/1618.
- J. H. Oort. Possible Interpretations of the High-Velocity Clouds. *BAN*, 18:421–438, 1966.

- E. Papastergis, A. M. Martin, R. Giovanelli, and M. P. Haynes. The Velocity Width Function of Galaxies from the 40% ALFALFA Survey: Shedding Light on the Cold Dark Matter Overabundance Problem. *ApJ*, 739:38, September 2011. doi: 10.1088/0004-637X/739/1/38.
- J. Peñarrubia, A. J. Benson, M. G. Walker, G. Gilmore, A. W. McConnachie, and L. Mayer. The impact of dark matter cusps and cores on the satellite galaxy population around spiral galaxies. *MNRAS*, 406:1290–1305, August 2010. doi: 10.1111/j.1365-2966.2010.16762.x.
- J. E. G. Peek, M. E. Putman, and J. Sommer-Larsen. Ongoing Galactic Accretion: Simulations and Observations of Condensed Gas in Hot Halos. *ApJ*, 674:227–236, February 2008. doi: 10.1086/524374.
- J. E. G. Peek, C. Heiles, K. A. Douglas, M.-Y. Lee, J. Grcevich, S. Stanimirović, M. E. Putman, E. J. Korpela, S. J. Gibson, A. Begum, D. Saul, T. Robishaw, and M. Krčo. The GALFA-HI Survey: Data Release 1. *ApJS*, 194:20, June 2011. doi: 10.1088/0067-0049/194/2/20.
- D. J. Pisano, D. G. Barnes, B. K. Gibson, L. Staveley-Smith, K. C. Freeman, and V. A. Kilborn. Where Are the High-Velocity Clouds in Local Group Analogs? *ApJL*, 610:L17–L20, July 2004. doi: 10.1086/423239.
- D. J. Pisano, D. G. Barnes, B. K. Gibson, L. Staveley-Smith, K. C. Freeman, and V. A. Kilborn. An HI Survey of Six Local Group Analogs. I. Survey Description and the Search for High-Velocity Clouds. *ApJ*, 662:959–968, June 2007. doi: 10.1086/517986.
- D. J. Pisano, D. G. Barnes, L. Staveley-Smith, B. K. Gibson, V. A. Kilborn, and K. C. Freeman. An HI Survey of Six Local Group Analogs. II. HI Properties of Group Galaxies. *ApJS*, 197:28, December 2011. doi: 10.1088/0067-0049/197/2/28.
- M. E. Putman, V. de Heij, L. Staveley-Smith, R. Braun, K. C. Freeman, B. K. Gibson, W. B. Burton, D. G. Barnes, G. D. Banks, R. Bhathal, W. J. G. de Blok, P. J. Boyce, M. J. Disney, M. J. Drinkwater, R. D. Ekers, P. A. Henning, H. Jerjen, V. A. Kilborn, P. M. Knezek, B. Koribalski, D. F. Malin, M. Marquarding, R. F. Minchin, J. R. Mould, T. Oosterloo, R. M. Price, S. D. Ryder, E. M. Sadler, I. Stewart, F. Stootman, R. L. Webster, and A. E. Wright. HIPASS High-Velocity Clouds: Properties of the Compact and Extended Populations. *AJ*, 123:873–891, February 2002. doi: 10.1086/338088.
- K. L. Rhode, J. J. Salzer, N. C. Haurberg, A. Van Sistine, M. D. Young, M. P. Haynes, R. Giovanelli, J. M. Cannon, E. D. Skillman, K. B. W. McQuinn, and E. A. K. Adams. ALFALFA Discovery of the Nearby Gas-rich Dwarf Galaxy Leo P. II. Optical Imaging Observations. *AJ*, 145:149, June 2013. doi: 10.1088/0004-6256/145/6/149.

- J. W. Rich, W. J. G. de Blok, T. J. Cornwell, E. Brinks, F. Walter, I. Bagetakos, and R. C. Kennicutt, Jr. Multi-Scale CLEAN: A Comparison of its Performance Against Classical CLEAN on Galaxies Using THINGS. *AJ*, 136:2897, December 2008. doi: 10.1088/0004-6256/136/6/2897.
- M. Ricotti. Late gas accretion on to primordial minihaloes: a model for Leo T, dark galaxies and extragalactic high-velocity clouds. *MNRAS*, 392:L45–L49, January 2009. doi: 10.1111/j.1745-3933.2008.00586.x.
- M. Rocha, A. H. G. Peter, and J. Bullock. Infall times for Milky Way satellites from their present-day kinematics. *MNRAS*, 425:231–244, September 2012. doi: 10.1111/j.1365-2966.2012.21432.x.
- E. V. Ryan-Weber, A. Begum, T. Oosterloo, S. Pal, M. J. Irwin, V. Belokurov, N. W. Evans, and D. B. Zucker. The Local Group dwarf Leo T: HI on the brink of star formation. *MNRAS*, 384:535–540, February 2008. doi: 10.1111/j.1365-2966.2007.12734.x.
- A. Saintonge. The Arecibo Legacy Fast ALFA Survey. IV. Strategies for Signal Identification and Survey Catalog Reliability. *AJ*, 133:2087–2096, May 2007. doi: 10.1086/513515.
- D. J. Sand, J. Strader, B. Willman, D. Zaritsky, B. McLeod, N. Caldwell, A. Seth, and E. Olszewski. Tidal Signatures in the Faintest Milky Way Satellites: The Detailed Properties of Leo V, Pisces II, and Canes Venatici II. *ApJ*, 756:79, September 2012. doi: 10.1088/0004-637X/756/1/79.
- J. Sand, D. The Structure and Star Formation of the New Milky Way Satellites and Beyond. In *American Astronomical Society Meeting Abstracts*, volume 222 of *American Astronomical Society Meeting Abstracts*, page 314.01, June 2013.
- D. R. Saul, J. E. G. Peek, J. Grcevich, M. E. Putman, K. A. Douglas, E. J. Korpela, S. Stanimirović, C. Heiles, S. J. Gibson, M. Lee, A. Begum, A. R. H. Brown, B. Burkhart, E. T. Hamden, N. M. Pingel, and S. Tonnesen. The GALFA-H I Compact Cloud Catalog. *ApJ*, 758:44, October 2012. doi: 10.1088/0004-637X/758/1/44.
- J. M. Schombert, S. S. McGaugh, and J. A. Eder. Gas Mass Fractions and the Evolution of Low Surface Brightness Dwarf Galaxies. *AJ*, 121:2420–2430, May 2001. doi: 10.1086/320398.
- G. S. Shostak and E. D. Skillman. Neutral hydrogen observations of the irregular galaxy IC 10. *A&A*, 214:33–42, April 1989.
- J. D. Simon and M. Geha. The Kinematics of the Ultra-faint Milky Way Satellites: Solving the Missing Satellite Problem. *ApJ*, 670:313–331, November 2007. doi: 10.1086/521816.

- E. D. Skillman, R. Terlevich, P. J. Teuben, and H. van Woerden. H I synthesis observations of the dwarf irregular galaxy Sextans A. *A&A*, 198:33–42, June 1988.
- E. D. Skillman, J. J. Salzer, D. A. Berg, R. W. Pogge, N. C. Haurberg, J. M. Cannon, E. Aver, K. A. Olive, R. Giovanelli, M. P. Haynes, E. A. K. Adams, K. B. W. McQuinn, and K. L. Rhode. ALFALFA Discovery of the nearby Gas-rich Dwarf Galaxy Leo P. III. An Extremely Metal Deficient Galaxy. *AJ*, 146:3, July 2013. doi: 10.1088/0004-6256/146/1/3.
- S. Stanimirović, M. Putman, C. Heiles, J. E. G. Peek, P. F. Goldsmith, B.-C. Koo, M. Krčo, J.-J. Lee, J. Mock, E. Muller, J. D. Pandian, A. Parsons, Y. Tang, and D. Werthimer. First Results from the Arecibo Galactic H I Survey: The Disk/Halo Interface Region in the Outer Galaxy. *ApJ*, 653:1210–1225, December 2006. doi: 10.1086/508800.
- A. Sternberg, C. F. McKee, and M. G. Wolfire. Atomic Hydrogen Gas in Dark Matter Minihalos and the Compact High-Velocity Clouds. *ApJS*, 143:419–453, December 2002. doi: 10.1086/343032.
- P. B. Stetson. DAOPHOT - A computer program for crowded-field stellar photometry. *PASP*, 99:191–222, March 1987. doi: 10.1086/131977.
- J. M. Stil, F. J. Lockman, A. R. Taylor, J. M. Dickey, D. W. Kavars, P. G. Martin, T. A. Rothwell, A. I. Boothroyd, and N. M. McClure-Griffiths. Compact H I Clouds at High Forbidden Velocities in the Inner Galaxy. *ApJ*, 637:366–379, January 2006. doi: 10.1086/498347.
- L. E. Strigari, J. S. Bullock, M. Kaplinghat, J. D. Simon, M. Geha, B. Willman, and M. G. Walker. A common mass scale for satellite galaxies of the Milky Way. *Nature*, 454:1096–1097, August 2008. doi: 10.1038/nature07222.
- M. Teyssier, K. V. Johnston, and M. Kuhlen. Identifying Local Group field galaxies that have interacted with the Milky Way. *MNRAS*, 426:1808–1818, November 2012. doi: 10.1111/j.1365-2966.2012.21793.x.
- R. Teyssier, A. Pontzen, Y. Dubois, and J. I. Read. Cusp-core transformations in dwarf galaxies: observational predictions. *MNRAS*, 429:3068–3078, March 2013. doi: 10.1093/mnras/sts563.
- E. J. Tollerud, J. S. Bullock, L. E. Strigari, and B. Willman. Hundreds of Milky Way Satellites? Luminosity Bias in the Satellite Luminosity Function. *ApJ*, 688: 277–289, November 2008. doi: 10.1086/592102.
- D. Tseliakhovich and C. Hirata. Relative velocity of dark matter and baryonic fluids and the formation of the first structures. *PhRvD*, 82(8):083520, October 2010. doi: 10.1103/PhysRevD.82.083520.

- P. G. van Dokkum. Cosmic-Ray Rejection by Laplacian Edge Detection. *PASP*, 113:1420–1427, November 2001. doi: 10.1086/323894.
- B. P. Wakker and H. van Woerden. Distribution and origin of high-velocity clouds. III - Clouds, complexes and populations. *A&A*, 250:509–532, October 1991.
- B. P. Wakker and H. van Woerden. High-Velocity Clouds. *ARA&A*, 35:217–266, 1997. doi: 10.1146/annurev.astro.35.1.217.
- S. M. Walsh, B. Willman, and H. Jerjen. The Invisibles: A Detection Algorithm to Trace the Faintest Milky Way Satellites. *AJ*, 137:450–469, January 2009. doi: 10.1088/0004-6256/137/1/450.
- T. Westmeier, R. Braun, and D. Thilker. Westerbork H I observations of high-velocity clouds near M 31 and M 33. *A&A*, 436:101–115, June 2005a. doi: 10.1051/0004-6361:20052867.
- T. Westmeier, C. Brüns, and J. Kerp. Effelsberg H I observations of compact high-velocity clouds. *A&A*, 432:937–953, March 2005b. doi: 10.1051/0004-6361:20041246.
- B. Willman. In Pursuit of the Least Luminous Galaxies. *Advances in Astronomy*, 2010:285454, 2010. doi: 10.1155/2010/285454.
- A. Zolotov, A. M. Brooks, B. Willman, F. Governato, A. Pontzen, C. Christensen, A. Dekel, T. Quinn, S. Shen, and J. Wadsley. Baryons Matter: Why Luminous Satellite Galaxies have Reduced Central Masses. *ApJ*, 761:71, December 2012. doi: 10.1088/0004-637X/761/1/71.
- M. A. Zwaan. A targeted survey for H i clouds in galaxy groups. *MNRAS*, 325: 1142–1148, August 2001. doi: 10.1046/j.1365-8711.2001.04514.x.